

This Page Is Insert  
and is not a part of

**BEST AVAILABLE**

Defective images within this document are accurate representation  
the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

**IMAGES ARE BEST AVAILABLE COPY.**

As rescanning documents *will not* correct images,  
please do not report the images to the  
Image Problem Mailbox.

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:36 ; Search time 399.68 seconds  
(without alignments)  
3.057 Million cell updates/sec

Title: US-09-446-543a-74  
Perfect score: 47  
Sequence: 1 SRXHXSMEXR 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	91.5	31	AAW31384	Rat type G protein
2	43	91.5	31	AAW97233	Rat type 11gand po
3	43	91.5	31	AAW67614	Rat 19p2 11gand.
4	43	91.5	31	AAW95173	Murine pituitary-d
5	43	91.5	31	AAW95174	Murine pituitary-d
6	43	91.5	31	AAW10355	Rat oxytocin secre
7	43	91.5	31	AAW87504	Rat prolactin-rele
8	43	91.5	31	AAW49292	19p2 11gand peptid
9	43	91.5	31	AAW62524	Rat CRH releasin
10	43	91.5	31	AAW90993	Prolactin releasin
11	43	91.5	32	AAW31385	Rat type G protein

12	43	91.5	32	21	AAW10356	Rat oxytocin secre
13	43	91.5	32	22	AAW62525	Rat CRH releasin
14	43	91.5	33	18	AAW31386	Rat type G protein
15	43	91.5	33	21	AAW10357	Rat oxytocin secre
16	43	91.5	33	22	AAW62526	Rat CRH releasin
17	43	91.5	32	20	AAW95172	Murine pituitary-d
18	43	91.5	33	18	AAW31383	Rat type G protein
19	43	91.5	33	20	AAW97225	Rat type 11gand po
20	43	91.5	33	21	AAW10354	Rat oxytocin secre
21	43	91.5	83	22	AAW62523	Rat CRH releasin
22	42	89.4	15	18	AAW31389	Synthetic 11gand 1
23	42	89.4	15	20	AAW97229	N-terminal 11gand
24	42	89.4	29	18	AAW31369	Bovine G protein-c
25	42	89.4	29	20	AAW95164	Bovine pituitary-d
26	42	89.4	30	21	AAW49299	19p2 11gand peptid
27	42	89.4	31	18	AAW1391	Human type G prote
28	42	89.4	31	18	AAW31371	Bovine G protein-c
29	42	89.4	31	20	AAW97235	Human type 11gand
30	42	89.4	31	20	AAW97218	Bovine pituitary-d
31	42	89.4	31	20	AAW87613	Bovine 19p2 11gand
32	42	89.4	31	20	AAW67615	Human 19p2 11gand
33	42	89.4	31	20	AAW95188	Bovine pituitary-d
34	42	89.4	31	21	AAW10347	Bovine oxytocin se
35	42	89.4	31	21	AAW10362	Human oxytocin sec
36	42	89.4	31	21	AAW49290	19p2 11gand peptid
37	42	89.4	31	21	AAW49291	19p2 11gand peptid
38	42	89.4	31	21	AAW49298	Bovine CRH releas1
39	42	89.4	31	22	AAW62516	Human CRH releasin
40	42	89.4	31	22	AAW62531	Prolactin releasin
41	42	89.4	31	22	AAW80991	Prolactin releasin
42	42	89.4	31	22	AAW90995	Prolactin releasin
43	42	89.4	31	22	AAW87370	bPrp31 peptid, u
44	42	89.4	32	18	AAW31392	Human type G prote
45	42	89.4	32	18	AAW31372	Bovine G protein-c

## ALIGNMENTS

RESULT 1	AAW31384	standard; Peptide: 31 AA.
ID	AAW31384	
AC	AAW31384	
XX	06-APR-1998	(first entry)
DE	Rat type G protein-coupled receptor 11gand fragment 1.	
XX	G protein-coupled receptor; 11gand binding; pharmaceutical;	
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;	
KW	therapeutic agent.	
OS	Rat sp.	
XX	WO9724436-A2.	
PN	10-JUL-1997.	
PD		
XX	26-DEC-1996;	96W0-JP03821.
PF		
XX	18-SEP-1996;	96JP-0246573.
PR	28-DEC-1995;	95JP-034371.
PR	15-MAR-1996;	96JP-0059419.
PR	12-AUG-1996;	96JP-0211805.
XX	(TAKE ) TAKEDA CHEM IND LTD.	
FA	Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;	
PI	Kawamata Y, Kitada C;	
PI	WPI, 1997-363672/33.	
XX	N-PSDB; AAW02421.	
DR		

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS  
 XX Claim 2; Page 179; 258pp: English.

CC This sequence represents a peptide fragment from a novel rat type  
 CC ligand polypeptide corresponding to amino acid residues 22 to 52 of the  
 CC sequence represented in AAW31383 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC hyperlipidemia, hypercholesterolemia, hyper- and polypylagia,  
 CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligogalactia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.

SQ Sequence 31 AA;

Query Match 91.5%; Score 43; DB 18; Length 31;

Best Local Similarity 72.7%; Pred. No. 0.026; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
 II I III I  
 Db 1 strahsmetr 11

RESULT 2

AAW97233 ID AAW97233 standard; peptide; 31 AA.

AC AAW97233;

DT 06-MAY-1999 (first entry)

DE Rat type ligand polypeptide fragment.

XX Rat type ligand: modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypocoarctinism; gonocyst cogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumor; amenorrhea; galactorrhea;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chlari-Frommel syndrome; Argon-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysosperma;  
 KW contraceptive; placental function; chorioncarcinoma; hydralid mole;  
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.

OS Ratus sp.

PN WO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumors, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3; Page 153; 241pp: English.

CC The present sequence represents a rat type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypocoarctinism, gonocyst cogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumor, amenorrhea, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chlari-Frommel syndrome, lymphoma, Sheehan syndrome or dysosperma.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC chorioncarcinoma, hydralid mole, interruption mole, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

SQ Sequence 31 AA;

Query Match 91.5%; Score 43; DB 20; Length 31;

Best Local Similarity 72.7%; Pred. No. 0.026; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
 II I III I  
 Db 1 strahsmetr 11

RESULT 3

AAW87614 ID AAW87614 standard; peptide; 31 AA.

AC AAW87614;

DT 29-MAR-1999 (first entry)

DE Rat 19P2 ligand.

XX 19P2 ligand: G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; rat; dementia; breast cancer;  
 KW therapy.

OS Ratus sp.

PN EP887417-A2.

PD 30-DEC-1998.

PF 25-JUN-1998; 98EP-0111725.

PR 27-JUN-1997; 97JP-0172118.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Moriya T, Nishimura O, Suenaga M, Tanaka Y;

DR WPI; 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
 PT of a fusion protein, useful for preventing and treating dementia,  
 PT breast cancer, renal failure and autoimmune disease  
 XX  
 PS Claim 5; Page 34; 56pp: English.

XX This is the amino acid sequence of the rat pituitary G  
 CC protein-coupled receptor ligand 19p2L. A method suitable for  
 CC commercial high-level production of 19p2L comprises expressing  
 CC the ligand in host cells as a recombinant fusion protein e.g. with  
 CC human basic fibroblast growth factor (see AY83794-95) that has  
 CC been modified to include an N-terminal cysteine residue. The  
 CC ligand is released from the fusion by cyanylation followed by  
 CC ammonolysis. 19p2L has prolactin secretion-stimulating and (at  
 CC high doses) prolactin secretion-inhibiting properties. It can be  
 CC used in the treatment and prevention of various diseases including:  
 CC senile dementia, cerebrovascular dementia, and dementia associated  
 CC with: genealogical disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease, Pick's disease, Huntington's disease), infectious diseases  
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
 CC toxemia (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
 CC intoxication by drugs, metal and organic compounds), tumorigenic  
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
 CC subarachnoidal hemorrhage, and other types of dementia, depression,  
 CC hyperactive child syndrome (microcephalopathy) and disturbance of  
 CC consciousness. It is also useful for prevention and treatment of  
 CC diseases associated with prolactin hypo and hypersecretion  
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,  
 CC breast cancer, infertility, impotence and autoimmune disease  
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,  
 CC osteoporosis, menopausal syndrome and renal failure (hypersecretion  
 CC disorders). The 19p2 polypeptide/amide is also useful as a test  
 CC reagent for study of the prolactin secretory function or as a  
 CC lactagogue in mammalian farm animals.

XX Sequence 31 AA:

Query Match 91.5%; Score 43; DB 20; Length 31;

Best Local Similarity 72.7%; Pred. No. 0.026; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

YY 1 SRXHXSMEXR 11  
 || |||||  
 Db 1 strahqmetr 11

RESULT 4

AAW95173 standard; peptide: 31 AA.

XX AAW95173:

XX 10-MAR-1999 (first entry)

XX Murine pituitary-derived ligand mature polypeptide sequence.

XX Pituitary-derived ligand polypeptide: G-protein coupled orphan receptor;  
 KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KM gene therapy; transgenic animal.

XX Mus sp.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hinuma S;

XX WPI; 1999-009423/01.  
 XX New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening

XX Disclosure; Page 134; 206pp; English.

XX This represents the matured murine pituitary-derived ligand polypeptide  
 CC sequence. The polypeptide is a ligand for the G-protein coupled orphan  
 CC receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with  
 CC a vector containing the ligand polypeptide encoding DNA are used to  
 CC produce a recombinant ligand polypeptide. The ligand polypeptide, and its  
 CC fragments, modulate function of the pituitary, central nervous system,  
 CC pancreas and other tissues and can be used to screen for agents that  
 CC modulate binding of the polypeptide to the receptor; to quantify the  
 CC amount of receptor in a sample and to raise antibodies. They may also be  
 CC used therapeutically, e.g. to treat senile dementia; Alzheimer's,  
 CC Parkinson's or Huntington's diseases; Creutzfeldt-Jakob disease; poisoning  
 CC by heavy metals or drugs; diabetes; schizophrenia; disorders of growth  
 CC hormone secretion; cancer; rheumatoid arthritis, epilepsy and many  
 CC others, also to improve post-operative nutritional status and as  
 CC vasopressor. Transgenic animals carrying the ligand polypeptide encoding  
 CC DNA or its mutein are used to study the function of the polypeptide-  
 CC expressing genes, as models of disease, for drug screening and as source  
 CC of cell lines. The ligand polypeptide DNA is used as a source of probes  
 CC and primers; to identify related sequences; in receptor-binding assays;  
 CC for production of Ab and antisera; in drug development; for gene therapy  
 CC and to develop transgenic animals.

XX Sequence 31 AA:

Query Match 91.5%; Score 43; DB 20; Length 31;

Best Local Similarity 72.7%; Pred. No. 0.026; Mismatches 0; Indels 0; Gaps 0;

Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

YY 1 SRXHXSMEXR 11  
 || |||||  
 Db 1 strahqmetr 11

RESULT 5

AAW95174 standard; Protein: 31 AA.

XX AAW95174:

XX 10-MAR-1999 (first entry)

XX Murine pituitary-derived ligand polypeptide antigenic epitope.

XX Pituitary-derived ligand polypeptide: G-protein coupled orphan receptor;  
 KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KM gene therapy; transgenic animal; epitope.

XX Mus sp.

XX WO9849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE ) TAKEDA CHEM IND LTD.

Pt	Fukusumi S., Hinuma S;
Df	WPI: 1999-009423/01.
Pt	New polypeptide ligand for orphan G protein coupled receptors - used
Pt	for treating disorders of central nervous system, pituitary and
Pt	pancreas, and for drug screening
Pt	
Pt	Disclosure; Page 26; 206pp; English.
Pt	
Pt	The invention relates to a murine pituitary-derived ligand polypeptide
Pt	which is a ligand for the G-protein coupled orphan receptor designated
Pt	GPR10 (human) or UMR-1 (rat). Cells transformed with a vector containing
Pt	the ligand polypeptide encoding DNA are used to produce a recombinant
Pt	ligand polypeptide. The ligand polypeptide, and its fragments, modulate
Pt	function of the pituitary, central nervous system, pancreas and other
Pt	tissues and can be used to screen for agents that modulate binding of
Pt	the polypeptide to the receptor; to quantify the amount of receptor in a
Pt	sample and to raise antibodies. They may also be used therapeutically,
Pt	e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's
Pt	diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;
Pt	diabetes; schizophrenia; disorders of growth hormone secretion; cancer;
Pt	rheumatoid arthritis, epilepsy and many others, also to improve post-
Pt	operative nutritional status and as vasopressor. Transgenic animals
Pt	carrying the ligand polypeptide encoding DNA or its mutagen are used to
Pt	study the function of the polypeptide-expressing genes, as models of
Pt	disease, for drug screening and as source of cell lines. The ligand
Pt	polypeptide DNA is used as a source of probes and primers; to identify
Pt	related sequences; in receptor-binding assays; for production of Ab and
Pt	antisera; in drug development; for gene therapy and to develop
Pt	transgenic animals. Sequences AAM95174 to AAM95178 represent antigenic
Pt	epitopes which can be used for the preparation of anti-ligand polypeptide
Pt	antibody.
Pt	
Pt	Sequence    31 AA:
Pt	
Pt	SQ
Pt	XX
Pt	XX
Pt	Query Match                      91.5%;    Score 43;    DB 20;    Length 31;
Pt	Best Local Similarity         72.7%;    Pred. No. 0.026;
Pt	Matches    8; Conservative        0; Mismatches    3; Indels    0; Gaps    0
Oy	1 SRXKHXSMEXR 11
Oy	
Oy	1 strahqsmetr 11
Oy	
Db	
RESULT 6	.
AAB10355	
AA10355 standard; peptide; 31 AA.	
AA10355;	
AA10355;	
24-NOV-2000 (first entry)	
Rat oxytocin secretion promoting peptide SEQ ID NO: 18.	
Rat oxytocin secretion promoting peptide SEQ ID NO: 18.	
Rat oxytocin secretion promoter; G protein-coupled receptor protein;	
treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;	
cassarean section; artificial fertilization; galactostasis; goat; pig;	
veterinary medicine; milk production.	
Rattus sp.	
WO200038704-A1.	
06-JUL-2000.	
22-DEC-1999; 99WO-JP07199.	
25-DEC-1998; 98JP-0369585.	
(TAKE ) TAKEDA CHEM IND LTD.	

PI	Matsumoto H, Kitada C, Hinuma S;
XX	
DR	WPI: 2000-452298/39.
XX	
PT	Physiologically-active polypeptide recognized as ligand by G
PR	protein-coupled receptor protein, for promoting secretion of oxytocin,
PT	as drugs for diseases relating to oxytocin secretion and in veterinary
XX	medicine -
PS	Claim 3; Page 57; 72pp; Japanese.
CC	
CC	This invention describes a novel oxytocin secretion-regulating agent
CC	which contains a ligand peptide or its salt for the G protein-coupled
CC	receptor protein. It is useful in the form of drugs for ameliorating,
CC	preventing and treating diseases relating to oxytocin secretion e.g.,
CC	weak pains and atonic bleeding, before and after expulsion of placenta,
CC	uterine recovery failure, caesarean section, stoppage of artificial
CC	fertilization or galactostasis and is also applicable in veterinary
CC	medicine for promoting milk production in cow, goat and pig. This
CC	sequence represents a rat peptide which acts as an oxytocin secretion
XX	promoter.
SQ	
Sequence	31 AA;
OY	1 SRXHXSMEXR 11             db 1 srahgsmetr 11
RESULT 7	
AAy87504	
ID	AAy87504 standard; protein; 31 AA.
AC	AAy87504;
XX	
DT	18-JUL-2000 (first entry)
DE	Rat prolactin-releasing peptide, PRP.
XX	
KM	Prolactin-releasing peptide; PRP; GPR10; G protein-coupled receptor;
KM	feeding behaviour; food intake; modulation; antagonist; anorectic;
KW	obesity; agonist; cachexia.
XX	
OS	Rattus sp.
FH	
FT	Key Location/Qualifiers
FT	Modified-site 31
FT	/note= "C-terminal amide"
PN	WO200017641-A1.
PD	30-MAR-2000.
PF	22-SEP-1999; 99WO-US21243.
PR	22-SEP-1998; 98US-0101380.
PR	14-OCT-1998; 98US-0172353.
PA	(MILL-) MILLENNIUM PHARM INC.
PL	Stricker-Kongrad A, Gu W;
DR	WPI: 2000-303231/26.
PT	Identifying modulators of body weight by a combination of a cell-free
PT	or cell-based assay to identify modulators of GPR10, followed by an in
PT	vivo assay for the compounds effect on e.g. feeding behavior -
XX	

PS Example 2; Page 61; 82pp; English.  
XX  
CC The invention relates to a method for identifying compounds useful for  
CC modulating body weight. The method comprises cell-free and/or cell-based  
CC assays that identify compounds which bind to and/or activate or inhibit  
CC the activity of GPR10, a G protein-coupled receptor. These assays are  
CC then followed by an in vivo assay of the effect of the compound on  
CC feeding behaviour, body weight or metabolic rate in a mammal. Proactin-  
CC releasing peptide (PrRP; AAY67504) is a ligand of GPR10. Binding of PrRP  
CC to GPR10 stimulates a signal transduction cascade, which results in an  
CC increase in food intake. Compounds identified using the method of the  
CC invention are useful for the modulation of body weight. Antagonists of  
CC GPR10 can be used to treat obesity, while GPR10 agonists can be used to  
CC treat cachexia. The present sequence represents rat PrRP.  
XX  
SQ Sequence 31 AA;  
  
Query Match 91.5%; Score 43; DB 21; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.026;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 SRXHXSMEXR 11  
Db 1 srhxsmetr 11  
  
RESULT 8  
AAY49292  
ID AAY49292 standard; peptide: 31 AA.  
XX  
AC AAY49292;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE 19P2 ligand peptide fragment.  
XX  
KM Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
OS Rattus sp.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 31  
FT /note="C-terminal amide"  
XX  
PN WO9960112-A1.  
XX  
PD 25-NOV-1999.  
XX  
PF 20-MAY-1999; 99WO-JP02650.  
XX  
PR 21-MAY-1998; 98JP-0140293.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-039381/03.  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
PS Disclosure; Page 26; 73pp; Japanese.  
XX  
CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.

CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.  
XX  
SQ Sequence 31 AA;  
  
Query Match 91.5%; Score 43; DB 21; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.026;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 SRXHXSMEXR 11  
Db 1 srhxsmetr 11  
  
RESULT 9  
AAG62524  
ID AAG62524 standard; peptide: 31 AA.  
XX  
AC AAG62524;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Rat CRH releasing protein related peptide SEQ ID NO: 18.  
XX  
KM Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
XX Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Rattus sp.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
XX  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
XX  
DR WPI; 2001-355552/37.  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Claim 3; Page 69; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 31 AA;  
  
Query Match 91.5%; Score 43; DB 22; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.026;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
  
Qy 1 SRXHXSMEXR 11  
Db 1 srhxsmetr 11

RESULT 10  
ID AAB90993 standard; Peptide; 31 AA.  
XX  
AC AAB90993;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Proactin releasing peptide SEQ ID NO:167.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KM blood component; modification; succinimide; maleimide group; amino;  
KM hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
PS  
XX  
PS Disclosure: Page 244; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (II) and a  
CC reactive group (III) (e.g. succinimide and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 31 AA;

Query Match 91.5%; Score 43; DB 22; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.026; 3; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
II | | | | |  
Db 1 srhghsmetr 11

RESULT 11  
ID AAW31385 standard; Peptide; 32 AA.  
XX

AC AAW31385;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Rat type G protein-coupled receptor ligand fragment 2.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KM therapeutic agent.  
XX  
OS Rat sp.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI; 1997-363672/23.  
XX  
PT N-PSDB; AAV02422.  
XX  
PS  
XX  
PS Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 179; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 22 to 53 of the  
CC sequence represented in AAW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. This ligand could have specific applications as a  
CC propylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC Turner's syndrome, neurosis, diabetes, cancer, pancreatitis, renal disease,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosacchara. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 32 AA;

Query Match 91.5%; Score 43; DB 18; Length 32;  
Best Local Similarity 72.7%; Pred. No. 0.026; 3; Indels 0; Gaps 0;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
II | | | | |  
Db 1 srhghsmetr 11

RESULT 12  
ID AAB10356 standard; Peptide; 32 AA.  
XX  
AC AAB10356;

XX 24-NOV-2000 (first entry)  
 DT Rat oxytocin secretion promoting peptide SEQ ID NO: 19.  
 XX  
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 19.  
 XX  
 KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Rattus sp.  
 XX WO200038704-A1.  
 PN  
 XX  
 XX 06-JUL-2000.  
 PD  
 XX 22-DEC-1999; 99WO-JP07199.  
 PF  
 XX 25-DEC-1998; 98JP-0369585.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PA Matsumoto H, Kitada C, Hinuma S;  
 PI WPI; 2000-452298/39.  
 DR  
 XX Physiologically-active polypeptide recognized as ligand by G  
 XX protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Disclosure; Page 57; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a rat peptide which acts as an oxytocin secretion  
 CC promoter.  
 CC  
 XX  
 SQ Sequence 32 AA:  
 XX  
 QY 1 SRXHXSMEXR 11  
 || | || | |  
 Db 1 srhxhsmetr 11

Query Match 91.5%; Score 43; DB 21; Length 32;  
 Best Local Similarity 72.7%; Pred. NO. 0.026; Indels 0; Gaps 0;  
 Matches 8; Conservative 0; Mismatches 3;

RESULT 13  
 AAG62525  
 ID AAG62525 standard; peptide; 32 AA.  
 XX  
 AC AAG62525;  
 XX  
 DT 24-AUG-2001 (first entry)  
 XX  
 DE Rat CRH releasing protein related peptide SEQ ID NO: 19.  
 XX  
 KW Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.  
 XX  
 OS Rattus sp.  
 XX WO200135984-A1.  
 PN

XX 25-MAY-2001.  
 PD  
 XX  
 XX 17-NOV-2000; 2000WO-JP08119.  
 PF  
 XX 18-NOV-1999; 99JP-0327900.  
 PR 26-SEP-2000; 2000JP-0297073.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PA Kitada C, Matsumoto H, Hinuma S;  
 PI WPI; 2001-355552/37.  
 DR  
 XX  
 XX Use of G protein receptor ligand or peptide for controlling  
 PT corticotropin releasing hormone secretion -  
 PT  
 PS Disclosure; Page 69; 90pp; Japanese.  
 XX  
 CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention.  
 CC  
 XX  
 SQ Sequence 32 AA:  
 XX  
 QY 1 SRXHXSMEXR 11  
 || | || | |  
 Db 1 srhxhsmetr 11

Query Match 91.5%; Score 43; DB 22; Length 32;  
 Best Local Similarity 72.7%; Pred. NO. 0.026; Indels 0; Gaps 0;  
 Matches 8; Conservative 0; Mismatches 3;

RESULT 14  
 AAW31386  
 ID AAW31386 standard; peptide; 33 AA.  
 XX  
 AC AAW31386;  
 XX  
 DT 06-APR-1998 (first entry)  
 XX  
 DE Rat type G protein-coupled receptor ligand fragment 3.  
 XX  
 KW G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Rat sp.  
 XX  
 PN WO9724436-A2.  
 XX  
 XX 10-JUL-1997.  
 PD  
 XX 26-DEC-1996; 96WO-JP03821.  
 PF  
 XX 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-034371.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
 PI Kawamata Y, Kitada C;

XX WPI; 1997-363672/33.  
 DR N-PSDB; AAV02423.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 179-180; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel rat type  
 CC ligand polypeptide corresponding to amino acid residues 22 to 54 of the  
 CC sequence represented in AAW31383 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and hypothyria,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC Turner's syndrome, diabetes, cancer, pancreatitis, renal disease,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosacchara. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 33 AA;  
 XX  
 Query Match 91.5%; Score 43; DB 18; Length 33;  
 Best Local Similarity 72.7%; Pred. No. 0.027;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 SRXHXSMEXR 11  
 ||| ||||| |  
 Db 1 srahqsmetr 11  
 XX  
 RESULT 15  
 AAB10357  
 ID AAB10357 standard; peptide; 33 AA.  
 XX  
 AC AAB10357;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 20.  
 XX  
 KW Rat; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Rattus sp.  
 OS  
 PN WO200038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX  
 DR WPI; 2000-452298/39.  
 XX  
 PT Physiologically-active polypeptide recognized as ligand by G

PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Disclosure; Page 58; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a rat peptide which acts as an oxytocin secretion  
 CC promoter.  
 XX  
 SQ Sequence 33 AA;  
 XX  
 Query Match 91.5%; Score 43; DB 21; Length 33;  
 Best Local Similarity 72.7%; Pred. No. 0.027;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 OY 1 SRXHXSMEXR 11  
 ||| ||||| |  
 Db 1 srahqsmetr 11  
 XX

Search completed: September 13, 2002, 09:18:36  
 Job time: 501 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 13, 2002, 09:06:40 : Search time 30.03 seconds  
(without alignments)  
122.059 Million cell updates/sec

Title: SEQ74-FUSED-TO-SEQ73  
Perfect score: 156  
Sequence: 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRFX 33

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	94.2	31	18 AAW31391	Human type G prote
2	147	94.2	31	18 AAW31384	Rat type G protein
3	147	94.2	31	20 AAW97233	Rat type ligand po
4	147	94.2	31	20 AAW97235	Human type ligand
5	147	94.2	31	20 AAW87614	Rat 19P2 ligand.
6	147	94.2	31	20 AAW87615	Human 19P2 ligand.
7	147	94.2	31	20 AAW95173	Murine pituitary-d
8	147	94.2	31	20 AAW95174	Murine pituitary-d
9	147	94.2	31	21 AAB10355	Rat oxytocin secre
10	147	94.2	31	21 AAB10362	Human oxytocin sec
11	147	94.2	31	21 AAW87504	Rat prolactin-rele

12	147	94.2	31	21 AAY49291	19P2 ligand peptid
13	147	94.2	31	21 AAY49292	19P2 ligand peptid
14	147	94.2	31	22 AAG62524	Rat CRH releasin
15	147	94.2	31	22 AAG62531	Human CRH releasin
16	147	94.2	31	22 AAB90991	Proactin releasin
17	147	94.2	31	22 AAB90995	Proactin releasin
18	147	94.2	31	22 AAB90995	Proactin releasin
19	147	94.2	32	18 AAW31392	Human type G prote
20	147	94.2	32	18 AAW31385	Rat type G protein
21	147	94.2	32	21 AAB10356	Rat oxytocin secre
22	147	94.2	32	21 AAB10363	Human oxytocin sec
23	147	94.2	32	22 AAG62525	Rat CRH releasin
24	147	94.2	32	22 AAG62532	Human CRH releasin
25	147	94.2	33	18 AAW31393	Human type G prote
26	147	94.2	33	18 AAW31386	Rat type G protein
27	147	94.2	33	21 AAB10357	Rat oxytocin secre
28	147	94.2	33	21 AAB10364	Human oxytocin sec
29	147	94.2	33	22 AAG62526	Rat CRH releasin
30	147	94.2	33	22 AAG62533	Human CRH releasin
31	147	94.2	82	20 AAW95172	Murine pituitary-d
32	147	94.2	83	18 AAW31383	Rat type G protein
33	147	94.2	83	20 AAW97225	Rat type ligand po
34	147	94.2	83	21 AAB10354	Rat oxytocin secre
35	147	94.2	83	22 AAG62523	Rat CRH releasin
36	147	94.2	87	18 AAW31390	Human type G prote
37	147	94.2	87	20 AAW97226	Human type ligand
38	147	94.2	87	21 AAB10361	Human oxytocin sec
39	147	94.2	87	22 AAG62530	Human CRH releasin
40	146	93.6	31	18 AAW31371	Bovine G protein-G
41	146	93.6	31	20 AAW97218	Bovine pituitary-d
42	146	93.6	31	20 AAW87613	Bovine 19P2 ligand
43	146	93.6	31	20 AAW95188	Bovine pituitary-d
44	146	93.6	31	21 AAB10347	Bovine oxytocin se
45	146	93.6	31	21 AAY49290	19P2 ligand peptid

## ALIGNMENTS

RESULT 1

AAW31391

ID AAW31391 standard: Peptide: 31 AA.

XX

AC AAW31391:

XX

DT 06-APR-1998 (first entry)

XX

DE Human type G protein-coupled receptor ligand fragment 1.

XX

KW G protein-coupled receptor; ligand binding; pharmaceutical;

KW modulator; pituitary; central nervous system; pancreas; prophylactic;

KW therapeutic agent.

XX

OS Homo sapiens.

XX

PN WO9724436-A2.

XX

PD 10-JUL-1997.

XX

PF 26-DEC-1996; 96WO-JP03821.

XX

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343373.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX

PA (TAKE ) TAKEDA CHEM IND LTD.

XX

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,

PI Kawamata Y, Kikada C;

XX

DR WPI: 1997-363672/33.

DR N-PSDB: AAW02426.



modulating prolactin  
in menopausal  
pregnancy

[illegible]

standard; peptide; 31 AA.

AAW97235;  
XX 06-MAY-1999 (first entry)  
wide fragment

DT	U6-mw	Human type ligand polypeptide ligand	
DE	XX	Human type ligand; modulation; prolactin secretion; gonecyst caecogenesis	
XX	XX	Rat type ligand; receptor; GPCR; hypocoarctanism; lactation; autolism	
XX	XX	G protein-coupled receptor; eutryoid; hypometabolism; lactation; autolism	
XX	XX	menopausal syndrome; brain tumour; amenorrhea; galactorrhea;	
XX	XX	pituitary adenomatosis; infertility; impotence; Argon-del Castillo syndrome;	
XX	XX	prolactinoma; infertile; lymphoma; Sheehan syndrome; hydatid mole;	
XX	XX	acromegaly; Chlath syndrome; chorioncarcinoma; abnormal saccharometabolism	
XX	XX	corruptive; abortion; unfertility fetus; abnormal	
XX	XX	lipid metabolism; oxytocin.	
XX	XX	abnormal lipid metabolism; oxytocin.	
XX	XX	Homo sapiens.	
OS	XX	Homo sapiens.	
XX	XX	W09858962-A1.	
PN	XX	W09858962-A1.	
XX	XX	30-DEC-1998.	
PD	XX	98WO-JP02765.	
XX	XX	22-JUN-1998;	
PF	XX	97JP-0165437.	
XX	XX	23-JUN-1997;	
PR	XX	97JP-0165437.	
XX	XX	(TAKE ) TAKEDA CHEM IND LTD.	
PA	XX	Fujii R, Hinuma S, Kawamata Y, Matsumoto H;	
XX	XX	Fujii R, Hinuma S, Kawamata Y, Matsumoto H;	
XX	XX	WPI. 1999-105614/09.	
DR	XX	WPI. 1999-105614/09.	

XX use of G protein-coupled receptor  
PT secretion or placental function; disease or abnormality  
PT syndrome, tumours, autoimmune disease  
PT syndrome, tumours, autoimmune disease

159: 241pp; English.

type ligand fragment. It describes

XX syndrome, tumour.  
PS Claim 3; page 159; 2A1pp; English.  
XX The present sequence represents a human type ligand fragment. It  
XX is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a GPCR  
CC as used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which can be used for  
CC ligand polypeptides or promoting prolactin secretion, menopause  
CC protein. The agents for promoting prolactin secretion, menopause  
CC treating or hypothyroidism, they can be used for promoting  
CC treatment or hypothyroidism and as used for treating or preventing  
CC syndrome, eddyloid or hypometabolism. They can be used for  
CC lactation in a domestic mammal and as used for treating or preventing  
CC inhibition prolactin secretion, brain tumour, amenorrhoea, galactorrhea,  
CC inhibiting adenomatous, breast tumour, amenorrhoea, galactorrhea,  
CC pituitary adenoma, infertility, impotence, argonza del Castil syndrome,  
CC prolectinoma, chlamydia, lymphoma, Sheehan syndrome or dyszoospermia.  
CC acromegaly, infertility, impotence, argonza del Castil syndrome,  
CC Forbes-Albright syndrome, also be used for treating or preventing  
CC The inhibitory prolactin function, abortion, unthriftiness, oestrogen  
CC modulating prolactin function, abortion, unthriftiness, oestrogen  
CC chorioncarcinoma, hydatid mole, irruption mole, or oxytocin.  
CC abnormal saccharometaformism, abnormal lipidmetabolism or oxytocin.  
XX Sequence 31 AA:  
S0

Score 147, DB 20: Length 31:  
aa-17: aa-17: 0: Gaps  
0

	94.28;	Score	No. 2
Query Match	83.9%;	Pred.	Mismatch
Best Local	Conservative	0;	
Matches	26;		
OY	1 SRXHXSMEXRPDPINPMWXXRGIRPVGE	31	
Db	1 sthrismetrlpdpinpwysarglrpygrt	31	

RESULT 5  
nontride; 31 AA

AAW87614	standard; repl
ID	AAW87614
XX	
AC	AAW87614;
	(first entry)

DT 43 Rat 19P2 ligand.  
XX XX Rat 19P2 ligand.  
DE XX Rat 19P2 ligand.  
XX XX Rat 19P2 ligand.  
KM KM Rat 19P2 ligand.  
KM KM Rat 19P2 ligand.

XX	Rattus sp.
OS	
XX	EP887417-A2.
PN	
XX	30-DEC-1998.
PD	
XX	98EP-0111725.
XX	25-JUN-1998:
PR	
XX	27-JUN-1997: 97JP-0172118.
XX	
XX	(TAKE ) TAKEEDA CHEM IND LTD.
XX	
PA	
XX	
XX	Moriya T, Nishimura O, Suenaga M, Tanaka Y;
PI	
XX	WPI: 1999-047884/05.
DR	
XX	Producing a 19p2 plutitary G protein receptor ligand - by cleavage
PT	of a fusion protein useful for preventing and autoimmune disease
PT	Breast cancer, renal failure and autoimmune disease
XX	
PS	Claim 5: page 34: 56pp: English.
XX	

Query Match  
2st Floor

standard; Peptide; 31 AA.  
 x  
 PA

fusion a 19p2.

nd - by cleavage  
ting dementia,

breast cancer, renal failure and autoimmune disease

claim 5, Page 35, 56pp, English.

This is the amino acid sequence of the human pituitary G protein-coupled receptor ligand 192L. A method suitable for the ligand in high-level production of 192L comprises culturing human basic fibroblast as a recombinant fusion protein expressing ligand is released to include an N-terminal (see AA83196-97) that has aminoamylase. 192L has prolactin secretion-stimulating followed by used in the treatment and prevention of various diseases. (at disease) dementia, cerebrovascular dementia, and dementias including with: neurological disorders (e.g. Alzheimer's disease associated with: genetic/epigenetic disorders (e.g. Huntington's disease), infectious diseases (e.g. hypothyroidism, vitamin B12 deficiency, disease or hypercholesterolemia, brain tumor), traumatic diseases or subarachnoid hemorrhage, and other types of dementia (e.g. chronic conduction). It is also useful for prevention and treatment of diseases associated with prolactin (hyperprolactinemia, pituitary adenoma, breast cancer, infertility, hyperprolactinemia, and disturbance of (hypersecretion disorders), and senile and autoimmune disease disorders). The 192L polypeptide and renal failure (hyposecretion reagent for study of the prolactin/and is also useful as a test lactagogue in mammalian farm animals.

Sequence 31 AA:

Query Match	94.28	Score 147	DB 20	Length 31
Best Local Similarity	83.98	Pred. No. 2, 3e-17		
Matches 26	Conservative 0	Mismatches 5	Indels 0	Gaps 0
1	SRKHSHNEKTPDINPAPYXXRGIRPVGR	31		
1				
1	srthshmeirtpdinpawyasgiltrpvgrf	31		
1				
SPLIT 7				

AAW5173			
ID	AAW5173	standard; peptide; 31 AA.	
AC	AAW5173;		
XX			
DT	10-MAR-1999	(first entry)	
XX			
DE	Murine pituitary-derived ligand mature polypeptide sequence.		
XX			
KW	pituitary-derived ligand polypeptide; G-protein coupled orphan receptor.		
KW	GPCR0; UHR-1; modulator; pituitary; central nervous system; pancreas;		
KW	tissue; screen; therapeutic; binding; senile dementia; ligand; murine;		
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;		
KW	Creutzfeldt-Jakob disease; poliovirus; schizophrenia; growth hormone;		
KW	secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;		
KW	gene therapy; transgenic animal.		
OS	Mus sp.		
XX			
PN	W09849295-A1.		
XX			
PD	05-NOV-1998.		
XX			
PF	27-APR-1998;	98WO-JP01923.	
XX			
XX			
PR	28-APR-1997,	97JP-0109974	
XX			

PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fukusumi S, Hinuma S;  
 XX  
 DR WPI: 1999-009423/01.  
 XX  
 PT New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX  
 PS Disclosure: Page 134; 206pp; English.  
 XX  
 CC This represents the matured murine pituitary-derived ligand polypeptide  
 CC sequence. The polypeptide is a ligand for the G-protein coupled orphan  
 CC receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with  
 CC a vector containing the ligand polypeptide encoding DNA are used to  
 CC produce a recombinant ligand polypeptide. The ligand polypeptide, and its  
 CC fragments, modulate function of the pituitary, central nervous system,  
 CC pancreas and other tissues and can be used to screen for agents that  
 CC modulate binding of the polypeptide to the receptor; to quantify the  
 CC amount of receptor in a sample and to raise antibodies. They may also be  
 CC used therapeutically, e.g. to treat senile dementia; Alzheimer's,  
 CC Parkinson's or Huntington's diseases; Creutzfeldt-Jakob disease; poisoning  
 CC by heavy metals or drugs; diabetes; schizophrenia; disorders of growth  
 CC hormone secretion; cancer; rheumatoid arthritis, epilepsy and many  
 CC others, also to improve post-operative nutritional status and as  
 CC vasopressor. Transgenic animals carrying the ligand polypeptide encoding  
 CC DNA or its mutain are used to study the function of the polypeptide-  
 CC expressing genes, as models of disease, for drug screening and as source  
 CC of cell lines. The ligand polypeptide DNA is used as a source of probes  
 CC and primers; to identify related sequences; in receptor-binding assays;  
 CC for production of Ab and antisera; in drug development; for gene therapy  
 CC and to develop transgenic animals.  
 CC  
 SQ Sequence 31 AA:  
 XX  
 XX  
 Query Match 94.2%; Score 147; DB 20; Length 31;  
 Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 SRXHXSMEXRTPDINPAWXXRGIRPVGRF 31  
 ||| ||||| ||||| ||||| ||||| |||||  
 Db 1 srahghsmetrpdpinpawytcgrtprvgrf 31  
 RESULT 8  
 AAW95174  
 ID AAW95174 standard; Protein; 31 AA.  
 XX  
 AC AAW95174;  
 XX  
 DT 10-MAR-1999 (first entry)  
 XX  
 DE Murine pituitary-derived ligand polypeptide antigenic epitope.  
 XX  
 DE Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KW gene therapy; transgenic animal; epitope.  
 XX  
 OS Mus sp.  
 XX  
 PN WO9849295-A1.  
 XX  
 PD 05-NOV-1998.  
 XX  
 PF 27-APR-1998; 98WO-JP01923.  
 XX  
 PR 28-APR-1997; 97JP-0109974.

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fukusumi S, Hinuma S;  
 XX  
 DR WPI: 1999-009423/01.  
 XX  
 PT New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX  
 PS Disclosure: Page 26; 206pp; English.  
 XX  
 CC The invention relates to a murine pituitary-derived ligand polypeptide  
 CC which is a ligand for the G-protein coupled orphan receptor designated  
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
 CC the ligand polypeptide encoding DNA are used to produce a recombinant  
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
 CC function of the pituitary, central nervous system, pancreas and other  
 CC tissues and can be used to screen for agents that modulate binding of  
 CC the polypeptide to the receptor; to quantify the amount of receptor in a  
 CC sample and to raise antibodies. They may also be used therapeutically,  
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
 CC operative nutritional status and as vasopressor. Transgenic animals  
 CC carrying the ligand polypeptide encoding DNA or its mutain are used to  
 CC study the function of the polypeptide-expressing genes, as models of  
 CC disease, for drug screening and as source of cell lines. The ligand  
 CC polypeptide DNA is used as a source of probes and primers; to identify  
 CC related sequences; in receptor-binding assays; for production of Ab and  
 CC antisera; in drug development; for gene therapy and to develop  
 CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
 CC epitopes which can be used for the preparation of anti-ligand polypeptide  
 CC antibody.  
 CC  
 SQ Sequence 31 AA:  
 XX  
 XX  
 Query Match 94.2%; Score 147; DB 20; Length 31;  
 Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 SRXHXSMEXRTPDINPAWXXRGIRPVGRF 31  
 ||| ||||| ||||| ||||| ||||| |||||  
 Db 1 srahghsmetrpdpinpawytcgrtprvgrf 31  
 RESULT 9  
 AAB10355  
 ID AAB10355 standard; peptide; 31 AA.  
 XX  
 AC AAB10355;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Rat oxytocin secretion promoting peptide SEQ ID NO: 18.  
 XX  
 DE Rat; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Rattus sp.  
 XX  
 PN WO200038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.

```

XX (TAKE ) TAKEDA CHEM IND LTD.
PA
XX Matsumoto H, Kitada C, Hinuma S;
PI
XX WPI; 2000-452298/39.
DR
XX Physiolegically-active polypeptide recognized as ligand by G
PT protein-coupled receptor protein, for promoting secretion of oxytocin,
PT as drugs for diseases relating to oxytocin secretion and in veterinary
PT medicine -
XX
PS Claim 3; Page 57; 72pp; Japanese.
XX
CC This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating,
CC preventing and treating diseases relating to oxytocin secretion e.g.,
CC weak pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a rat peptide which acts as an oxytocin secretion
CC promoter.
XX
SQ Sequence 31 AA;

Query Match          94.2%; Score 147; DB 21; Length 31;
Best Local Similarity 83.9%; Pred. No. 2.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31
   || |||| |||||||| ||||||||
Db 1 srhghsmetrlpdpinpawygtrgtrpvgrf 31

RESULT 10
AAB10362
ID AAB10362 standard; peptide; 31 AA.
XX
AC AAB10362;
XX
DT 24-NOV-2000 (first entry)
XX
DE Human oxytocin secretion promoting peptide SEQ ID NO: 32.
XX
KM Human: oxytocin secretion promoter; G protein-coupled receptor protein;
KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;
KM caesarean section; artificial fertilization; galactostasis; goat; pig;
KM veterinary medicine; milk production.
XX
OS Homo sapiens.
XX
PN WO200038704-A1.
XX
PD 06-JUL-2000.
XX
PF 22-DEC-1999; 99WO-JP07199.
XX
PR 25-DEC-1998; 98JP-0369585.
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Matsumoto H, Kitada C, Hinuma S;
XX
DR WPI; 2000-452298/39.
XX
PT Physiolegically-active polypeptide recognized as ligand by G
PT protein-coupled receptor protein, for promoting secretion of oxytocin,
PT as drugs for diseases relating to oxytocin secretion and in veterinary
PT medicine -
XX

```

```

PS Disclosure; Page 62; 72pp; Japanese.
XX
CC This invention describes a novel oxytocin secretion-regulating agent
CC which contains a ligand peptide or its salt for the G protein-coupled
CC receptor protein. It is useful in the form of drugs for ameliorating,
CC preventing and treating diseases relating to oxytocin secretion e.g.,
CC weak pains and atonic bleeding, before and after expulsion of placenta,
CC uterine recovery failure, caesarean section, stoppage of artificial
CC fertilization or galactostasis and is also applicable in veterinary
CC medicine for promoting milk production in cow, goat and pig. This
CC sequence represents a human peptide which acts as an oxytocin secretion
CC promoter.
XX
SQ Sequence 31 AA;

Query Match          94.2%; Score 147; DB 21; Length 31;
Best Local Similarity 83.9%; Pred. No. 2.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31
   || |||| |||||||| ||||||||
Db 1 srhghsmetrlpdpinpawygtrgtrpvgrf 31

RESULT 11
AAY87504
ID AAY87504 standard; protein; 31 AA.
XX
AC AAY87504;
XX
DT 18-JUL-2000 (first entry)
XX
DE Rat prolactin-releasing peptide, PrRP.
XX
KM Prolactin-releasing peptide; PrRP; GPR10; G protein-coupled receptor;
KM feeding behaviour; food intake; modulation; antagonist; anorectic;
KM obesity; agonist; cachexia.
XX
OS Rattus sp.
XX
FH Key Location/Qualifiers
FT Modified-site 31 /note="C-terminal amide"
XX
PN WO200017641-A1.
XX
PD 30-MAR-2000.
XX
PF 22-SEP-1999; 99WO-US21243.
XX
PR 22-SEP-1998; 98US-0101380.
PR 14-OCT-1998; 98US-0172353.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Stricker-Kongrad A, Gu W;
XX
DR WPI; 2000-303231/26.
XX
PT Identifying modulators of body weight by a combination of a cell-free
PT or cell-based assay to identify modulators of GPR10, followed by an in
PT vivo assay for the compounds effect on e.g. feeding behavior -
XX
PS Example 2; Page 61; 82pp; English.
XX
CC The invention relates to a method for identifying compounds useful for
CC modulating body weight. The method comprises cell-free and/or cell-based
CC assays that identify compounds which bind to and/or activate or inhibit
CC the activity of GPR10, a G protein-coupled receptor. These assays are
CC then followed by an in vivo assay of the effect of the compound on
CC feeding behaviour, body weight or metabolic rate in a mammal. Prolactin-
CC releasing peptide (PrRP; AAY87504) is a ligand of GPR10. Binding of PrRP

```

CC to GPR10 stimulates a signal transduction cascade, which results in an  
CC increase in food intake. Compounds identified using the method of the  
CC invention are useful for the modulation of body weight. Antagonists of  
CC GPR10 can be used to treat obesity, while GPR10 agonists can be used to  
CC treat cachexia. The present sequence represents rat PRRP.

XX  
SQ Sequence 31 AA:

Query Match 94.2%; Score 147; DB 21; Length 31;  
Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPYGRF 31  
||| ||||| ||||| ||||| ||||| |||||  
Db 1 srlnrhsmetrtpdinpawysrglrpygrf 31

RESULT 12

AAy49291  
ID AAy49291 standard; peptide: 31 AA.

AC AAY49291;

DT 22-FEB-2000 (first entry)

DE 19P2 ligand peptide fragment.

XX Monoclonal antibody; 19P2 ligand; diagnosis: prolactin secretion;

KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Homo sapiens.

XX Key Location/Qualifiers

FT Modified-site 31 /note="C-terminal amide"

XX WO9960112-A1.

PD 25-NOV-1999.

PF 20-MAY-1999; 99WO-JP02650.

PR 21-MAY-1998; 98JP-0140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kltada C, Hinuma S;

XX WPI: 2000-039381/03.

PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality

PS Disclosure; Page 26; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.

XX  
SQ Sequence 31 AA:

Query Match 94.2%; Score 147; DB 21; Length 31;  
Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPYGRF 31  
||| ||||| ||||| ||||| ||||| |||||  
Db 1 srlnrhsmetrtpdinpawysrglrpygrf 31

RESULT 13

AAy49292  
ID AAy49292 standard; peptide: 31 AA.

AC AAY49292;

DT 22-FEB-2000 (first entry)

DE 19P2 ligand peptide fragment.

XX Monoclonal antibody; 19P2 ligand; diagnosis: prolactin secretion;

KW pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Rattus sp.

XX Key Location/Qualifiers

FT Modified-site 31 /note="C-terminal amide"

XX WO9960112-A1.

PD 25-NOV-1999.

PF 20-MAY-1999; 99WO-JP02650.

PR 21-MAY-1998; 98JP-0140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kltada C, Hinuma S;

XX WPI: 2000-039381/03.

PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality

PS Disclosure; Page 26; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.

XX  
SQ Sequence 31 AA:

Query Match 94.2%; Score 147; DB 21; Length 31;  
Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPYGRF 31  
||| ||||| ||||| ||||| ||||| |||||  
Db 1 srlnrhsmetrtpdinpawysrglrpygrf 31

RESULT 14

AAy62524  
ID AAy62524 standard; peptide: 31 AA.

AC AAY62524;

DT 24-AUG-2001 (first entry)

XX

```

DE Rat CRH releasing protein related peptide SEQ ID NO: 18.
XX
XX Rat; corticotrophin releasing hormone; CRH; G protein receptor ligand;
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;
KW Addison's disease; adrenal gland hyperfunction; obesity.
XX
XX Rattus sp.
XX
XX WO200135984-A1.
XX
XX 25-MAY-2001.
XX
XX 17-NOV-2000; 2000WO-JP08119.
XX
XX 18-NOV-1999; 99JP-0327900.
XX
XX 26-SEP-2000; 2000JP-0297073.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Kitada C, Matsumoto H, Hinuma S;
XX
XX WPI: 2001-355552/37.
XX
XX Use of G protein receptor ligand or peptide for controlling
XX corticotropin releasing hormone secretion -
XX
XX Claim 3; Page 69; 90pp; Japanese.
XX
XX The present sequence describes a method of controlling the secretion of
XX corticotrophin releasing hormone (CRH), involving the use of a G protein
XX receptor ligand. This can be used to control the secretion of CRH and is
XX useful as an analgesic or for treating, preventing or ameliorating
XX diseases associated with CRH secretion such as hyperaldosteronism,
XX hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
XX disease (including boredom, nausea, pigmentation, hypogonadism, hair
XX loss, and hypotension), adrenal gland hypofunction and obesity. The
XX present sequence is a peptide used in the exemplification of the
XX invention.
XX
XX Sequence 31 AA;
XX
SQ

```

Query Match 94.2%; Score 147; DB 22; Length 31;  
Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31
   ||| ||||| ||||| ||||| ||||| |||||
DB 1 strhghsmetrlpdpinpwylgrgrpvgrf 31

```

RESULT 15  
AAG62531  
ID AAG62531 standard; peptide: 31 AA.  
XX  
XX AAG62531;  
XX  
XX 24-AUG-2001 (first entry)  
XX  
XX Human CRH releasing protein related peptide SEQ ID NO: 32.  
XX  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX Homo sapiens.  
XX  
XX WO200135984-A1.  
XX  
XX 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-JP08119.  
XX

```

PR 18-NOV-1999; 99JP-0327900.
PR 26-SEP-2000; 2000JP-0297073.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Kitada C, Matsumoto H, Hinuma S;
XX
XX WPI: 2001-355552/37.
XX
XX Use of G protein receptor ligand or peptide for controlling
XX corticotropin releasing hormone secretion -
XX
XX Claim 3; Page 73-74; 90pp; Japanese.
XX
XX The present sequence describes a method of controlling the secretion of
XX corticotrophin releasing hormone (CRH), involving the use of a G protein
XX receptor ligand. This can be used to control the secretion of CRH and is
XX useful as an analgesic or for treating, preventing or ameliorating
XX diseases associated with CRH secretion such as hyperaldosteronism,
XX hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
XX disease (including boredom, nausea, pigmentation, hypogonadism, hair
XX loss, and hypotension), adrenal gland hypofunction and obesity. The
XX present sequence is a peptide used in the exemplification of the
XX invention.
XX
XX Sequence 31 AA;
XX
SQ

```

Query Match 94.2%; Score 147; DB 22; Length 31;  
Best Local Similarity 83.9%; Pred. No. 2.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

```

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31
   ||| ||||| ||||| ||||| ||||| |||||
DB 1 strhghsmetrlpdpinpwylgrgrpvgrf 31

```

Search completed: September 13, 2002, 09:10:14  
Job time: 214 sec



us-09-446-543a-61.rag

SEA ID NO: 61  
 AC NO: AAW31391  
 Database: A-geneeq-032802

Mon Sep 16 09:13:08 2002

GenCore version 2.00  
 Copyright (c) 1993 - 2000 Gen Ltd.  
 Search time 399.68 Seconds  
 (without alignments)  
 8.615 Million cell updates/sec

OM protein - protein search, using  
 September 13, 1993  
 APAMYASHRGIRPVGRF 31

Run on: US, Gapext 0.5  
 111073796 residues

Title score: 747574

Perfect score: classifying chosen parameters:

Sequence: 0

Scoring fn: 2000000000

Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

A-geneeq-032802:\*

- 1: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	171	100.0	31	18	AAW31391	Human type G prote
2	171	100.0	31	20	AAW97235	Human type G prote
3	171	100.0	31	20	AAW87615	Human type G prote
4	171	100.0	31	21	AAW10362	Human type G prote
5	171	100.0	31	21	AAW49291	Human type G prote
6	171	100.0	31	22	AAW62531	Human type G prote
7	171	100.0	31	22	AAW90991	Human type G prote
8	171	100.0	31	22	AAW31392	Human type G prote
9	171	100.0	31	22	AAW10363	Human type G prote
10	171	100.0	31	22	AAW62532	Human type G prote
11	171	100.0	33	18	AAW31393	Human type G prote

12	171	100.0	33	21	AAW10364	Human oxytocin sec
13	171	100.0	33	22	AAW62533	Human CRH releasin
14	171	100.0	87	18	AAW31390	Human type G prote
15	171	100.0	87	20	AAW97226	Human type G prote
16	171	100.0	87	21	AAW10361	Human oxytocin sec
17	171	100.0	87	22	AAW62530	Human CRH releasin
18	171	100.0	30	22	AAW49299	Human type G prote
19	165	96.5	31	22	AAW90995	Bovine type G prote
20	162	94.7	31	22	AAW31371	Bovine type G prote
21	158	92.4	31	20	AAW87613	Bovine type G prote
22	158	92.4	31	20	AAW95188	Bovine type G prote
23	158	92.4	31	20	AAW95187	Bovine type G prote
24	158	92.4	31	21	AAW49298	Bovine type G prote
25	158	92.4	31	21	AAW49299	Bovine type G prote
26	158	92.4	31	22	AAW62516	Bovine type G prote
27	158	92.4	32	18	AAW31372	Bovine type G prote
28	158	92.4	32	21	AAW95189	Bovine type G prote
29	158	92.4	32	21	AAW10348	Bovine type G prote
30	158	92.4	32	22	AAW62517	Bovine type G prote
31	158	92.4	33	20	AAW95190	Bovine type G prote
32	158	92.4	33	20	AAW31373	Bovine type G prote
33	158	92.4	33	21	AAW10349	Bovine type G prote
34	158	92.4	33	21	AAW49297	Bovine type G prote
35	158	92.4	33	22	AAW62518	Bovine type G prote
36	158	92.4	33	22	AAW31382	Bovine type G prote
37	158	92.4	98	18	AAW31368	Bovine type G prote
38	158	92.4	98	20	AAW97224	Bovine type G prote
39	158	92.4	98	20	AAW95187	Bovine type G prote
40	158	92.4	98	21	AAW95187	Bovine type G prote
41	158	92.4	98	21	AAW10346	Bovine type G prote
42	158	92.4	98	21	AAW10346	Bovine type G prote
43	158	92.4	98	22	AAW62515	Bovine type G prote
44	158	92.4	98	22	AAW62522	Bovine type G prote
45	158	92.4	98	22	AAW62522	Bovine type G prote

## ALIGNMENTS

RESULT 1	AAW31391 standard; Peptide: 31 AA.
ID	AAW31391
AC	AAW31391
DT	06-APR-1998 (first entry)
DE	Human type G protein-coupled receptor ligand fragment 1.
XX	G protein-coupled receptor; ligand binding; pharmacological;
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;
KW	therapeutic agent.
OS	Homo sapiens.
XX	WO9724436-A2.
PN	10-JUL-1997.
XX	26-DEC-1996; 96WO-JP03821.
PF	18-SEP-1996; 96JP-0246573.
XX	28-DEC-1995; 95JP-0343371.
PR	15-MAR-1996; 96JP-0059419.
PR	12-AUG-1996; 96JP-0211805.
XX	(TAKE ) TAKEDA CHEM IND LTD.
PA	Fujita R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
PI	Kawamata Y, Kitada C;
PI	WPI: 1997-363672/33.
DR	N-PSDB; AAW02428.

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 184; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
 CC sequence represented in AAM97235 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator. A central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophyactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC Turner's syndrome, diabetes, cancer, pancreatitis, renal disease,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC oligosaccharia. Assays can also be developed to screen compounds and/or  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.

SO Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 18; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTSRHSWEIRTPDINPAWYASRGIRPYGRF 31  
 DB 1 srtrhsmeirtpdinpawyasrgirpygrf 31

RESULT 2  
 AAM97235  
 XX AAM97235 standard; peptide; 31 AA.  
 CC AAM97235;

06-MAY-1999 (first entry)

an type ligand polypeptide fragment.

KW The ligand; modulation; prolactin secretion;  
 XX coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;  
 OS syndrome; euthyroid; hypometabolism; lactation;  
 XX Homo nematosis; brain tumour; emmenopathy; autolimmune disease;  
 PN hypofertility; impotence; amenorrhea; galactorrhea;  
 XX MO9858962-Aldrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 PD 30-DEC-1998; Argonx-del Castillo syndrome;  
 XX clinical function; choriocarcinoma; hydatid mole;  
 PF unthriftly fetus; abnormal saccharometabolism;  
 XX ocytocia.  
 PR 22-JUN-1998;  
 XX 23-JUN-1998; 98WO-JP02;  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI 97JP-0165437.  
 XX Fujii R, Hinuma S, Kawamata Y, Matsumoto Y,

DR WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 XX syndrome, tumours, autoimmune disease or abnormal pregnancy  
 PS Claim 3; Page 159; 241pp; English.

CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chlari-Frommel syndrome, Argonx-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, irruption mole, abortion, unthriftly fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or ocytocia.

SO Sequence 31 AA;  
 Query Match 100.0%; Score 171; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTSRHSWEIRTPDINPAWYASRGIRPYGRF 31  
 DB 1 srtrhsmeirtpdinpawyasrgirpygrf 31

RESULT 3  
 AAM87615  
 XX AAM87615 standard; Peptide; 31 AA.  
 AC AAM87615;

29-MAR-1999 (first entry)

Human 19P2 ligand.

DE 19P2 ligand; G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; human; dementia; breast cancer;  
 KW therapy.  
 XX Homo sapiens.

OS Homo sapiens.  
 XX EP887417-A2.  
 PN 30-DEC-1998.

PD 25-JUN-1998; 98EP-011725.  
 PF 27-JUN-1997; 97JP-0172118.

PR (TAKE ) TAKEDA CHEM IND LTD.  
 XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;

PI WPI; 1999-047884/05.

DR Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
 PT of a fusion protein, useful for preventing and treating dementia,  
 PT breast cancer, renal failure and autoimmune disease  
 XX Claim 5, Page 35, 56pp, English

XX This is the amino acid sequence of the human pituitary G  
 CC protein-coupled receptor ligand 19p2L. A method suitable for  
 CC commercial high-level production of 19p2L comprises expressing  
 CC the ligand in host cells as a recombinant fusion protein e.g. with  
 CC human basic fibroblast growth factor (see AY83796-97) that has  
 CC been modified to include an N-terminal cysteine residue. The  
 CC ligand is released from the fusion by cyanilation followed by  
 CC ammonolysis. 19p2L has prolactin secretion-stimulating and (at  
 CC high doses) prolactin secretion-inhibiting properties. It can be  
 CC used in the treatment and prevention of various diseases including:  
 CC senile dementia, cerebrovascular dementia, and dementia associated  
 CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's  
 CC disease, Pick's disease, Huntington's disease), infectious diseases  
 CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
 CC toxemia (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
 CC intoxication by drugs, metal and organic compounds), tumourigenic  
 CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
 CC subarachnoidal haemorrhage, and other types of dementia, depression,  
 CC hyperactive child syndrome (microencephalopathy) and disturbance of  
 CC consciousness. It is also useful for prevention and treatment of  
 CC diseases associated with prolactin hypo and hypersecretion  
 CC respectively, including: hyperprolactinaemia, pituitary adenoma,  
 CC breast cancer, infertility, impotence and autoimmune disease  
 CC (hypersecretion disorders), and seminal vesicle hypoplasia,  
 CC osteoporosis, menopausal syndrome and renal failure (hypersecretion  
 CC disorders). The 19p2 polypeptide/amide is also useful as a test  
 CC reagent for study of the prolactin secretory function or as a  
 CC lactagogue in mammalian farm animals.  
 XX Sequence 31 AA:

Query Match 100.0%; Score 171; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 4, 1e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTRHSMETRTDPINPAMYASRGIRPVGRF 31  
 DE ||||||||||||||||||||  
 DB 1 srtrhsmetrtdpinpamyasrgirpvgrf 31

RESULT 4  
 AAB10362  
 ID AAB10362 standard; peptide; 31 AA.  
 XX  
 AC AAB10362;  
 XX

DT 24-NOV-2000 (first entry)  
 XX

DE Human oxytocin secretion promoting peptide SEQ ID NO: 32.  
 XX

KM Human: oxytocin secretion promoter; G protein-coupled receptor protein;  
 KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KM caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KM veterinary medicine; milk production.  
 XX

OS Homo sapiens.  
 XX

PN WO200038704-A1.  
 XX

PD 06-JUL-2000.  
 XX

PF 22-DEC-1999; 99WO-JP07199.  
 XX

PR 25-DEC-1998; 98JP-0369585.  
 XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX

PI Matsumoto H, Kitada C, Hinuma S;  
 XX

DR WPI; 2000-452298/39.  
 XX

PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Disclosure; Page 62; 72pp; Japanese.  
 XX

CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.,  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.  
 XX Sequence 31 AA:

Query Match 100.0%; Score 171; DB 21; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 4, 1e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTRHSMETRTDPINPAMYASRGIRPVGRF 31  
 DE ||||||||||||||||||||  
 DB 1 srtrhsmetrtdpinpamyasrgirpvgrf 31

RESULT 5  
 AAY49291  
 ID AAY49291 standard; peptide; 31 AA.  
 XX  
 AC AAY49291;  
 XX

DT 22-FEB-2000 (first entry)  
 XX

DE 19p2 ligand peptide fragment.  
 XX

KM Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
 KM pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX

OS Homo sapiens.  
 XX

FH Key Location/Qualifiers  
 FT Modified-site 31 /note="C-terminal amide"  
 FT  
 XX

PN WO960112-A1.  
 XX

PD 25-NOV-1999.  
 XX

PF 20-MAY-1999; 99WO-JP02650.  
 XX

PR 21-MAY-1998; 98JP-0140293.  
 XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX

PI Matsumoto H, Kitada C, Hinuma S;  
 XX

DR WPI; 2000-039381/03.  
 XX

PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 XX

PS Disclosure; Page 26; 73pp; Japanese.  
 XX

CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central

CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 31 AA;

Query Match 100.0%; Score 171; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRHRHSMETRPDINPAMYASRGIRPVGRF 31  
DB 1 srhrhsmetrlpdpinpawyasrgirpvgf 31

RESULT 6  
AAG62531  
ID AAG62531 standard; peptide; 31 AA.

AC AAG62531;

DF 24-AUG-2001 (first entry)

DE Human CRH releasing protein related peptide SEQ ID NO: 32.

XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX

OS Homo sapiens.

PN WO200135984-A1.

PD 25-MAY-2001.

PF 17-NOV-2000; 2000WO-JP08119.

PR 18-NOV-1999; 99JP-0327900.

PR 26-SEP-2000; 2000JP-0297073.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Kltada C, Matsumoto H, Hinuma S;

DR WPI; 2001-355552/37.

PT Use of G protein receptor ligand or peptide for controlling

PT corticotrophin releasing hormone secretion -

XX Claim 3; Page 73-74; 90pp; Japanese.

CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX

SQ Sequence 31 AA;

Query Match 100.0%; Score 171; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRHRHSMETRPDINPAMYASRGIRPVGRF 31  
DB 1 srhrhsmetrlpdpinpawyasrgirpvgf 31

DB 1 srhrhsmetrlpdpinpawyasrgirpvgf 31  
RESULT 7  
AAB90991  
ID AAB90991 standard; peptide; 31 AA.

AC AAB90991;

DF 22-JUN-2001 (first entry)

DE Prolactin releasing peptide SEQ ID NO:165.

XX Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimideyl; maleimide group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX

OS Homo sapiens.

PN WO200069900-A2.

PD 23-NOV-2000.

PF 17-MAY-2000; 2000WO-US13576.

PR 17-MAY-1999; 99US-0134406.

PR 10-SEP-1999; 99US-0153406.

PR 15-OCT-1999; 99US-0159783.

XX (CONJ-) CONJUCHEM INC.

XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;

DR WPI; 2001-112059/12.

PT Modifying and attaching therapeutic peptides to albumin prevents

PT peptidase degradation, useful for increasing length of in vivo activity

XX Disclosure; Page 244; 733pp; English.

CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimideyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specifically as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX

SQ Sequence 31 AA;

Query Match 100.0%; Score 171; DB 22; Length 31;  
Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRHRHSMETRPDINPAMYASRGIRPVGRF 31  
DB 1 srhrhsmetrlpdpinpawyasrgirpvgf 31

RESULT 8

AAW31392  
ID AAW31392 standard; Peptide: 32 AA.  
XX  
AC AAW31392;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 2.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059919.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB; AAV02429.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2: Page 185; 258pp; English.  
XX  
SQ This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 23 to 54 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharide. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 32 AA:

Query Match 100.0%; Score 171; DB 18; Length 32;  
Best local Similarity 100.0%; Pred. No. 4.2e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHRSMETRPDINPAMYASRGIRPVGRF 31  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 1 strhrsmetrlpdpinpawyasrgirpvgrf 31

RESULT 9  
AAB10363

AAB10363 standard; peptide: 32 AA.  
ID AAB10363;  
XX  
AC AAB10363;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 33.  
XX  
KW Human: oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI: 2000-452298/39.  
XX  
PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
PS Disclosure: Page 62; 72pp; Japanese.  
XX  
SQ This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
SQ Sequence 32 AA:

Query Match 100.0%; Score 171; DB 21; Length 32;  
Best local Similarity 100.0%; Pred. No. 4.2e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRRHRSMETRPDINPAMYASRGIRPVGRF 31  
XXXXXXXXXXXXXXXXXXXXXXXXXXXXXXXX  
Db 1 strhrsmetrlpdpinpawyasrgirpvgrf 31

RESULT 10  
AAG62532  
ID AAG62532 standard; peptide: 32 AA.  
XX  
AC AAG62532;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 33.  
XX  
KW Human: corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX

OS Homo sapiens.  
 PN WO200135984-A1.  
 XX  
 XX 25-MAY-2001.  
 PD  
 XX 17-NOV-2000; 2000WO-JP08119.  
 PF  
 XX 18-NOV-1999; 99JP-0327900.  
 PR 26-SEP-2000; 2000JP-0297073.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Kitada C, Matsumoto H, Hinuma S;  
 PI WPI; 2001-355552/37.  
 DR  
 XX Use of G protein receptor ligand or peptide for controlling  
 PT corticotropin releasing hormone secretion -  
 PS Disclosure; Page 74; 90pp; Japanese.  
 XX  
 XX The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention.  
 CC  
 SQ Sequence 32 AA;  
 XX  
 XX  
 Query Match 100.0%; Score 171; DB 22; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 4.2e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31  
 DB 1 srthrhsmetirpdpinpawyaasrgirpvgrf 31  
 ||||||||||||||||||||||||||||  
 RESULT 11  
 AAW31393  
 ID AAW31393 standard; peptide: 33 AA.  
 XX  
 AC AAW31393;  
 XX  
 DT 06-APR-1998 (first entry)  
 XX  
 DE Human type G protein-coupled receptor ligand fragment 3.  
 XX  
 KW G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 PN WO9724436-A2.  
 PD 10-JUL-1997.  
 XX  
 XX 26-DEC-1996; 96WO-JP03821.  
 PF  
 XX 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343711.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kitada C;  
 XX  
 XX WPI; 1997-363672/33.  
 DR N-PSDB; AAV02430.  
 DR  
 XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 185; 258pp; English.  
 XX  
 XX This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 55 of the  
 CC sequence represented in AAW31390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharide. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 CC  
 SQ Sequence 33 AA;  
 XX  
 XX  
 Query Match 100.0%; Score 171; DB 18; Length 33;  
 Best Local Similarity 100.0%; Pred. No. 4.4e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31  
 DB 1 srthrhsmetirpdpinpawyaasrgirpvgrf 31  
 ||||||||||||||||||||||||||||  
 RESULT 12  
 AAB10364  
 ID AAB10364 standard; peptide: 33 AA.  
 XX  
 AC AAB10364;  
 XX  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 34.  
 XX  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 PN WO200038704-A1.  
 PD 06-JUL-2000.  
 XX  
 XX 22-DEC-1999; 99WO-JP07199.  
 PF  
 XX 25-DEC-1998; 98JP-0369585.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX

DR WPI; 2000-452298/39.  
XX  
XX  
XX Physiolegically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
PS Disclosure; Page 62-63; 72pp; Japanese.  
XX  
XX This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
XX  
SQ Sequence 33 AA;  
  
Query Match 100.0%; Score 171; DB 21; Length 33;  
Best Local Similarity 100.0%; Pred. No. 4,4e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SRRHSHMEIRTPDINPAMYASRGIRPVGRF 31  
1 strthshmeirtpdinpamyasrgirpvgrf 31  
Db 1 strthshmeirtpdinpamyasrgirpvgrf 31  
  
RESULT 13  
AAG62533  
ID AAG62533 standard; peptide; 33 AA.  
XX  
XX AAG62533;  
XX  
XX 24-AUG-2001 (first entry)  
XX  
XX Human CRH releasing protein related peptide SEQ ID NO: 34.  
DE  
XX  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200135984-A1.  
XX  
XX 25-MAY-2001.  
PD  
XX  
XX 17-NOV-2000; 2000MO-JP08119.  
PF  
XX  
XX 18-NOV-1999; 99JP-0327900.  
PR  
XX  
XX 26-SEP-2000; 2000JP-0297073.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Kitada C, Matsumoto H, Hinuma S;  
PI  
XX  
XX WPI; 2001-355552/37.  
DR  
XX  
XX Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
PT  
XX  
XX Disclosure; Page 74; 90pp; Japanese.  
PS  
XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,

CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
XX  
SQ Sequence 33 AA;  
  
Query Match 100.0%; Score 171; DB 22; Length 33;  
Best Local Similarity 100.0%; Pred. No. 4,4e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 SRRHSHMEIRTPDINPAMYASRGIRPVGRF 31  
1 strthshmeirtpdinpamyasrgirpvgrf 31  
Db 1 strthshmeirtpdinpamyasrgirpvgrf 31  
  
RESULT 14  
AAW31390  
ID AAW31390 standard; Protein; 87 AA.  
XX  
XX  
XX AAW31390;  
AC  
XX  
XX 06-APR-1998 (first entry)  
DT  
XX  
XX Human type G protein-coupled receptor ligand encoded by PHO87.  
DE  
XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9724436-A2.  
XX  
XX 10-JUL-1997.  
PD  
XX  
XX 26-DEC-1996; 96WO-JP03821.  
PF  
XX  
XX 18-SEP-1996; 96JP-0246573.  
PR  
XX  
XX 28-DEC-1995; 95JP-0343371.  
PR  
XX  
XX 15-MAR-1996; 96JP-0059419.  
PR  
XX  
XX 12-AUG-1996; 96JP-0211805.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI  
XX  
XX Kawamata Y, Kitada C;  
PI  
XX  
XX WPI; 1997-363672/33.  
DR  
XX  
XX N-PDB; AAV02427.  
DR  
XX  
XX Claim 3; Page 183; 258pp; English.  
PS  
XX  
XX This sequence represents a novel human type ligand polypeptide encoded  
CC by PHO87 which is used in an assay to monitor ligand binding to the G  
CC protein-coupled receptor protein. Pharmaceutical compositions  
CC containing this ligand may be used as a pituitary function modulator, a  
CC central nervous system modulator or a pancreatic function modulator. This  
CC ligand could have specific applications as a prophylactic or therapeutic  
CC agent for dementia, depression, hyperkinetic syndrome, disturbance of  
CC consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone  
CC secretory disease, hyper- and polypnea, hypercholesterolemia,  
CC hyperglyceridaemia, hyperlipidaemia, hyperprolactinaemia, diabetes,  
CC cancer, pancreatitis, renal disease, Turner's syndrome, neurosis,  
CC rheumatoid arthritis, spinal injury, transient brain ischaemia,  
CC amyotrophic lateral sclerosis, acute myocardial infarction,  
CC spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis,  
CC osteoporosis, asthma, epilepsy, infertility and/or oligogalactia. Assays

CC can also be developed to screen compounds which are capable of altering  
 CC the binding activity of the ligand thus affecting activation of the G  
 CC protein-coupled receptor protein.

XX Sequence 87 AA;

Query Match 100.0%; Score 171; DB 18; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SRTHRSMETRPDINPANYASRGIRPVGRF 31  
 |||  
 DB 23 strhrsmetirpdpinpawyasrgirpvgf 53

# RESULT 15

AAW97226  
 ID AAW97226 standard; peptide; 87 AA.

XX AAW97226;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide.

KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cagogenesis;  
 KW menopausal syndrome; eutthyroid; hypometabolism; lactation; modulation;  
 KW pituitary adenomatosis; brain tumor; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
 KW interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia; prolactin secretion;  
 KW human type ligand.

XX Homo sapiens.

XX WO9858962-A1.

PN 30-DEC-1998.

PD 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI: 1999-105614/09.

DR N-PSDB; AAX15526.

PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy

PS Disclosure: Page 158; 241pp; English.

CC The present sequence represents a human type ligand polypeptide. The  
 CC specification describes an agent for modulating prolactin secretion  
 CC which comprises a ligand polypeptide or a salt, for a G protein-coupled  
 CC receptor (GPCR). The agents for promoting prolactin secretion  
 CC can be used for treating or preventing hypovarianism, gonocyst  
 CC cagogenesis, menopausal syndrome, eutthyroid or hypometabolism. They  
 CC can be used for promoting lactation in a domestic mammal and as an  
 CC aphrodisiac. The agents for inhibiting prolactin secretion can be used  
 CC for treating or preventing pituitary adenomatosis, brain tumor,  
 CC emmenopathy, autoimmune disease, prolactinoma, infertility, impotence,  
 CC amenorrhea, galactorrhea, acromegaly, Chiari-Frommel syndrome, Argonz-del  
 CC Castillo syndrome, Forbes-Albright syndrome, lymphoma, Sheehan syndrome  
 CC or dyszoospermia. The inhibitory agents can also be used as

CC contraceptives. The agents for modulating placental function can be used  
 CC for treating or preventing choriocarcinoma, hydatid mole, interruption mole,  
 CC abortion, unfertilized fetus, abnormal saccharometabolism, abnormal  
 CC lipidmetabolism or oxytocia.

XX Sequence 87 AA;

Query Match 100.0%; Score 171; DB 20; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 1.4e-18;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 SRTHRSMETRPDINPANYASRGIRPVGRF 31  
 |||  
 DB 23 strhrsmetirpdpinpawyasrgirpvgf 53

Search completed: September 13, 2002, 09:18:34  
 Job time: 499 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:34 ; Search time 399.68 seconds  
(without alignments)  
6.114 Million cell updates/sec

Title: US-09-446-543a-73  
Perfect score: 109  
Sequence: 1 TPDIINPAMYXXRGIRPVGFRFX 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	96.3	20	18	AAW31394 Human type G prote
2	105	96.3	20	20	AAW97236 Human type ligand
3	105	96.3	20	21	AAW10365 Human oxytocin sec
4	105	96.3	20	21	AAW49284 19P2 ligand peptid
5	105	96.3	20	22	AAW62534 Human CRH releasin
6	105	96.3	20	22	AAW90992 Proactin releasin
7	105	96.3	21	18	AAW31395 Human type G prote
8	105	96.3	21	21	AAW10366 Human oxytocin sec
9	105	96.3	21	22	AAW62535 Human CRH releasin
10	105	96.3	22	18	AAW31396 Human type G prote
11	105	96.3	22	21	AAW10367 Human oxytocin sec

12	105	96.3	22	22	AAW62536 Human CRH releasin
13	105	96.3	31	18	AAW31391 Human type G prote
14	105	96.3	31	20	AAW97235 Human type ligand
15	105	96.3	31	20	AAW87615 Human 19P2 ligand
16	105	96.3	31	21	AAW10362 Human oxytocin sec
17	105	96.3	31	21	AAW49281 19P2 ligand peptid
18	105	96.3	31	22	AAW62531 Human CRH releasin
19	105	96.3	31	22	AAW90991 Proactin releasin
20	105	96.3	31	22	AAW62532 Human type G prote
21	105	96.3	32	18	AAW31392 Human oxytocin sec
22	105	96.3	32	21	AAW10363 Human CRH releasin
23	105	96.3	32	22	AAW62533 Human type G prote
24	105	96.3	33	18	AAW31393 Human oxytocin sec
25	105	96.3	33	21	AAW10364 Human CRH releasin
26	105	96.3	33	22	AAW62533 Human CRH releasin
27	105	96.3	37	18	AAW31390 Human type G prote
28	105	96.3	37	20	AAW97226 Human type ligand
29	105	96.3	37	21	AAW10361 Human oxytocin sec
30	105	96.3	37	22	AAW62530 Human CRH releasin
31	104	95.4	20	18	AAW31387 Rat type G prote
32	104	95.4	20	18	AAW31374 Bovine pituitary-d
33	104	95.4	20	20	AAW97232 Rat type ligand po
34	104	95.4	20	20	AAW97234 Bovine pituitary-d
35	104	95.4	20	20	AAW95191 Murine pituitary-d
36	104	95.4	20	21	AAW95175 Bovine oxytocin se
37	104	95.4	20	21	AAW10350 Rat oxytocin secre
38	104	95.4	20	21	AAW10358 19P2 ligand peptid
39	104	95.4	20	21	AAW49301 19P2 ligand peptid
40	104	95.4	20	21	AAW49302 Bovine CRH releasl
41	104	95.4	20	22	AAW62519 Rat CRH releasin
42	104	95.4	20	22	AAW62527 Proactin releasin
43	104	95.4	20	22	AAW90994 Proactin releasin
44	104	95.4	20	22	AAW90996 Peptide PRNP20 fra
45	104	95.4	20	22	AAW6954

# ALIGNMENTS

RESULT 1  
AAW31394 standard; Peptide: 20 AA.  
XX  
AC AAW31394:  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
PI Kawamata Y, Kltada C;  
XX  
DR WPI, 1997-363672/33.  
DR N-PSDB; AAW02431.

SEE ID NO: 73  
AC NO: AAW31394  
Database: A-gene-seq-032802



XX PS Disclosure; Page 63; 72pp; Japanese.  
XX CC This invention describes a novel oxytocin secretion-regulating agent  
XX CC which contains a ligand peptide or its salt for the G protein-coupled  
XX CC receptor protein. It is useful in the form of drugs for ameliorating,  
XX CC preventing and treating diseases relating to oxytocin secretion e.g.  
XX CC weak pains and atonic bleeding, before and after expulsion of placenta,  
XX CC uterine recovery failure, caesarean section, stoppage of artificial  
XX CC fertilization or galactostasis and is also applicable in veterinary  
XX CC medicine for promoting milk production in cow, goat and pig. This  
XX CC sequence represents a human peptide which acts as an oxytocin secretion  
XX CC promoter.  
XX SQ Sequence 20 AA;  
  
Query Match 96.3%; Score 105; DB 21; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.6e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 TPDINPAWYXXRGIRPYGRF 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 tpdinpawyasrgirpygrf 20  
  
RESULT 4  
AAY49294  
ID AAY49294 standard; peptide; 20 AA.  
XX AAY49294;  
XX  
XX 22-FEB-2000 (first entry)  
XX  
XX 19P2 ligand peptide fragment.  
XX DE  
XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
XX KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX KM  
XX Homo sapiens.  
XX OS  
XX  
XX FH Key Location/Qualifiers  
XX FT Modified-site 20  
XX PN /note="C-terminal amide"  
XX  
XX WO960112-A1.  
XX  
XX 25-NOV-1999.  
XX PD  
XX 20-MAY-1999; 99WO-JP02650.  
XX PF  
XX 21-MAY-1998; 98JP-0140293.  
XX PR  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX PA  
XX Matsumoto H, Kitada C, Hinuma S;  
XX PI WPI; 2000-039381/03.  
XX DR  
XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
XX PT studying diseases related to ligand abnormality -  
XX  
XX PS Disclosure; Page 26; 73pp; Japanese.  
XX  
XX The invention provides a monoclonal antibody which has a specific  
XX CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
XX CC derivative. The antibodies can be used in diagnosis or to treat or  
XX CC prevent diseases associated with abnormality in the pituitary function  
XX CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
XX CC nervous regulatory mechanism, and pancreatic function regulatory  
XX CC mechanism. The antibody-based immunoassay can also be applied in  
XX CC clarifying the physiological functions of the ligand and its derivative.  
XX CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.

XX SQ Sequence 20 AA;  
  
Query Match 96.3%; Score 105; DB 21; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.6e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 TPDINPAWYXXRGIRPYGRF 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 tpdinpawyasrgirpygrf 20  
  
RESULT 5  
AAG62534  
ID AAG62534 standard; peptide; 20 AA.  
XX AAG62534;  
XX  
XX 24-AUG-2001 (first entry)  
XX  
XX Human CRH releasing protein related peptide SEQ ID NO: 35.  
XX DE  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
XX KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
XX KM Addison's disease; adrenal gland hyperfunction; obesity.  
XX OS  
XX Homo sapiens.  
XX  
XX PN WO200135984-A1.  
XX PD  
XX 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-JP08119.  
XX PF  
XX 18-NOV-1999; 99JP-0327900.  
XX PR  
XX 26-SEP-2000; 2000JP-0297073.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX PA  
XX Kitada C, Matsumoto H, Hinuma S;  
XX PI WPI; 2001-355552/37.  
XX DR  
XX Use of G protein receptor ligand or peptide for controlling  
XX PT corticotropin releasing hormone secretion -  
XX  
XX PS Claim 4; Page 75; 90pp; Japanese.  
XX  
XX The present sequence describes a method of controlling the secretion of  
XX CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
XX CC receptor ligand. This can be used to control the secretion of CRH and is  
XX CC useful as an analgesic or for treating, preventing or ameliorating  
XX CC diseases associated with CRH secretion such as hyperaldosteronism,  
XX CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
XX CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
XX CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
XX CC present sequence is a peptide used in the exemplification of the  
XX CC invention.  
XX  
XX SQ Sequence 20 AA;  
  
Query Match 96.3%; Score 105; DB 22; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2.6e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 TPDINPAWYXXRGIRPYGRF 20  
| | | | | | | | | | | | | | | | | | | | | |  
Db 1 tpdinpawyasrgirpygrf 20  
  
RESULT 6

AAB90992  
ID AAB90992 standard; Peptide: 20 AA.  
XX  
AC AAB90992;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Proactin releasing peptide SEQ ID NO:166.  
XX  
KM Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KM blood component; modification; succinimide; maleimide group; amino;  
KM hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI: 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents.  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
PS Disclosure: Page 244: 733pp: English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide) and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 20 AA:  
XX  
Query Match 96.3%; Score 105; DB 22; Length 20;  
Best Local Similarity 90.0%; Pred. No. 2,6e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 1 TPDINPAMYXXRGIRPVGRF 20

RESULT 7  
AAW31395  
ID AAW31395 standard; Peptide: 21 AA.  
XX  
AC AAW31395;  
XX

XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 5.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prolactin;  
KM therapeutic agent.  
XX  
OS Homo sapiens.  
OS  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 95JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR NPI: 1997-363672/33.  
DR N-PSDB: AAV02432.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2: Page 186; 258pp: English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, interstitial, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharide. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 21 AA:  
XX  
Query Match 96.3%; Score 105; DB 18; Length 21;  
Best Local Similarity 90.0%; Pred. No. 2,8e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 1 TPDINPAMYXXRGIRPVGRF 20

RESULT 8  
AAB10366  
ID AAB10366 standard; Peptide: 21 AA.  
XX  
AC AAB10366;  
XX

DT 24-NOV-2000 (first entry)  
 XX Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
 DE Human oxytocin secretion promoter; G protein-coupled receptor protein;  
 XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KM caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KM veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200038704-A1.  
 XX  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-JP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX WPI; 2000-452298/39.  
 DR  
 XX  
 PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Disclosure; Page 63; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.,  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.  
 CC  
 XX  
 SQ Sequence 21 AA;  
 XX  
 XX  
 Query Match 96.3%; Score 105; DB 21; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 2.8e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPDINPAWYXXRGIRPYGRF 20  
 ||||||| |||||||  
 Db 1 tpdinpawyasrglrpygrf 20

RESULT 9  
 AAG62535  
 ID AAG62535 standard; peptide; 21 AA.  
 XX  
 AC AAG62535;  
 XX  
 DT 24-AUG-2001 (first entry)  
 XX  
 DE Human CRH releasing protein related peptide SEQ ID NO: 36.  
 XX  
 KM Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KM analgesic; hyperaldosteronism; hypercortisolism; hypoadrenocorticism;  
 KM Addison's disease; adrenal gland hyperfunction; obesity.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200135984-A1.  
 XX

PD 25-MAY-2001.  
 XX  
 PF 17-NOV-2000; 2000WO-JP08119.  
 XX  
 PR 18-NOV-1999; 99JP-0327900.  
 PR 26-SEP-2000; 2000JP-0297073.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Kitada C, Matsumoto H, Hinuma S;  
 XX WPI; 2001-355552/37.  
 DR  
 XX  
 PT Use of G protein receptor ligand or peptide for controlling  
 PT corticotrophin releasing hormone secretion -  
 XX  
 PS Disclosure; Page 75; 90pp; Japanese.  
 XX  
 CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolism, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the  
 CC invention.  
 CC  
 XX  
 SQ Sequence 21 AA;  
 XX  
 XX  
 Query Match 96.3%; Score 105; DB 22; Length 21;  
 Best Local Similarity 90.0%; Pred. No. 2.8e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 TPDINPAWYXXRGIRPYGRF 20  
 ||||||| |||||||  
 Db 1 tpdinpawyasrglrpygrf 20

RESULT 10  
 AAW31396  
 ID AAW31396 standard; Peptide; 22 AA.  
 XX  
 AC AAW31396;  
 XX  
 DT 06-APR-1998 (first entry)  
 XX  
 DE Human type G protein-coupled receptor ligand fragment 6.  
 XX  
 KM G protein-coupled receptor; ligand binding; pharmaceutical;  
 KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KM therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9724436-A2.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 26-DEC-1996; 96WO-JP03821.  
 XX  
 PR 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343371.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujita R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kitada C;  
 XX

DR WPI: 1997-363672/33.  
 DR N-PSDB: AAV02433.  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2: Page 186; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
 CC sequence represented in AAW1390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidemia, hypercholesterolemia, hyperglycidaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligogalactia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 CC  
 SQ Sequence 22 AA:  
 OY 1 TPDINPAMYXXRGIRPVGRF 20  
 DB 1 tpdinpamyasrigrpvgrf 20  
 ||||||| |||||||  
 RESULT 11  
 AAB10367  
 ID AAB10367 standard; peptide: 22 AA.  
 AC AAB10367;  
 DT 24-NOV-2000 (first entry)  
 XX  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
 XX  
 KM Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KM caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KM veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200038704-A1.  
 PD 06-JUL-2000.  
 XX  
 PF 22-DEC-1999; 99WO-UP07199.  
 XX  
 PR 25-DEC-1998; 98JP-0369585.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kitada C, Hinuma S;  
 XX  
 DR WPI: 2000-452298/39.  
 PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,

PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 PS Disclosure; Page 64; 72pp; Japanese.  
 XX  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.  
 CC  
 SQ Sequence 22 AA:  
 OY 1 TPDINPAMYXXRGIRPVGRF 20  
 DB 1 tpdinpamyasrigrpvgrf 20  
 ||||||| |||||||  
 RESULT 12  
 AAG62536  
 ID AAG62536 standard; peptide: 22 AA.  
 AC AAG62536;  
 DT 24-AUG-2001 (first entry)  
 XX  
 DE Human CRH releasing protein related peptide SEQ ID NO: 37.  
 XX  
 KM Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KM analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
 KM Addison's disease; adrenal gland hyperfunction; obesity.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO200135984-A1.  
 PD 25-MAY-2001.  
 XX  
 PF 17-NOV-2000; 2000WO-UP08119.  
 XX  
 PR 18-NOV-1999; 99JP-0327900.  
 PR 26-SEP-2000; 2000JP-0297073.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Kitada C, Matsumoto H, Hinuma S;  
 XX  
 DR WPI: 2001-355552/37.  
 PT Use of G protein receptor ligand or peptide for controlling  
 PT corticotrophin releasing hormone secretion -  
 XX  
 PS Disclosure; Page 75; 90pp; Japanese.  
 XX  
 CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
 CC present sequence is a peptide used in the exemplification of the

CC Invention.  
XX  
SQ Sequence 22 AA;

Query Match 96.3%; Score 105; DB 22; Length 22;  
Best Local Similarity 90.0%; Pred. No. 2.9e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 TPDINPAWYXXRCIRPVGRF 20  
|||||||  
Db 1 tpdinpaawyasrgirpvgf 20

RESULT 13  
AAW31391  
ID AAW31391 standard; Peptide; 31 AA.  
XX  
AC AAW31391;

DT 06-APR-1998 (first entry)  
XX

DE Human type G protein-coupled receptor ligand fragment 1.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prolactin;  
XX therapeutic agent.

OS Homo sapiens.

XX WO9724436-A2.

PN 10-JUL-1997.

PD 26-DEC-1996; 96WO-JP03821.

PF 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX N-PSDB; AAW02428.

DR WPI: 1997-363672/33.

XX Claim 2; Page 184; 258pp; English.

This sequence represents a peptide fragment from a novel human type  
ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
sequence represented in AAW31390 and is used in an assay to monitor  
ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
compositions containing this ligand may be used as a pituitary function  
modulator, a central nervous system modulator or a pancreatic function  
modulator. This ligand could have specific applications as a  
prolactin, growth hormone secretory disease, hyper- and polypylagia,  
trauma, growth hormone secretory disease, hyper- and polypylagia,  
hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
acute myocardial infarction, interstitial, spinocerebellar degeneration,  
bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
oligodactylia. Assays can also be developed to screen compounds which are  
capable of altering the binding activity of the ligand affecting  
activation of the G protein-coupled receptor protein.

XX  
SQ Sequence 31 AA;

Query Match 96.3%; Score 105; DB 18; Length 31;  
Best Local Similarity 90.0%; Pred. No. 4.3e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 TPDINPAWYXXRCIRPVGRF 20  
|||||||  
Db 12 tpdinpaawyasrgirpvgf 31

RESULT 14  
AAW97235  
ID AAW97235 standard; peptide; 31 AA.  
XX  
AC AAW97235;

DT 06-MAY-1999 (first entry)  
XX

DE Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypocoarctinism; gonocyst cacosgenesis;  
KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; amenorrhoea; autolimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhoea; galactorrhoea;  
KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KM contraceptive; placental function; chorioncarcinoma; hydralid mole;  
KM interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KM abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WO968962-A1.

PN 30-DEC-1998.

PD 22-JUN-1996; 98WO-JP02765.

PF 23-JUN-1997; 97JP-0165437.

PR (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

PI WPI: 1999-105614/09.

DR Claim 3; Page 159; 241pp; English.

The present sequence represents a human type ligand fragment. It  
is used in the course of the invention. The specification describes  
an agent for modulating prolactin secretion which comprises a  
ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
protein. The agents for promoting prolactin secretion can be used for  
treating or preventing hypocoarctinism, gonocyst cacosgenesis, menopausal  
syndrome, euthyroid or hypometabolism. They can be used for promoting  
lactation in a domestic mammal and as an aphrodisiac. The agents for  
inhibiting prolactin secretion can be used for treating or preventing  
pituitary adenomatosis, brain tumour, emmenorrhoea, autolimmune disease,  
prolactinoma, infertility, impotence, amenorrhoea, galactorrhoea,  
acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
The inhibitory agents can also be used as contraceptives. The agents for  
modulating placental function can be used for treating or preventing  
chorioncarcinoma, hydralid mole, interruption mole, abortion, unfertilized fetus,  
abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 31 AA;  
SQ

Query Match  
Best Local Similarity 96.3%; Score 105; DB 20; Length 31;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYXXRGIRPVGRF 20  
|||||  
DB 12 tpdinpawyasrgirpvgrf 31

RESULT 15  
AAW87615  
ID AAW87615 standard; peptide: 31 AA.  
XX  
AC AAW87615;  
XX  
DT 29-MAR-1999 (first entry)  
XX  
DE Human 19P2 ligand.  
XX  
KW 19P2 ligand; G protein coupled receptor; pituitary;  
KM prolactin releasing peptide; human; dementia; breast cancer;  
KM therapy.  
XX  
OS Homo sapiens.  
XX  
PN EP887417-A2.  
XX  
PD 30-DEC-1998.  
XX  
PF 25-JUN-1998; 98EP-0111725.  
XX  
PR 27-JUN-1997; 97JP-0172118.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Moriya T, Nishimura O, Suenaga M, Tanaka Y;  
XX  
DR WPI; 1999-047884/05.  
XX  
PT Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
PT of a fusion protein, useful for preventing and treating dementia,  
PT breast cancer, renal failure and autoimmune disease  
XX  
PS Claim 5; Page 35; 56pp; English.  
XX

This is the amino acid sequence of the human pituitary G protein-coupled receptor ligand 19P2L. A method suitable for commercial high-level production of 19P2L comprises expressing the ligand in host cells as a recombinant fusion protein e.g. with human basic fibroblast growth factor (see AAV83796-97) that has been modified to include an N-terminal cysteine residue. The ligand is released from the fusion by cyanylation followed by ammonolysis. 19P2L has prolactin secretion-stimulating and (at high doses) prolactin secretion-inhibiting properties. It can be used in the treatment and prevention of various diseases including: senile dementia, cerebrovascular dementia, and dementia associated with: neurological disorders (e.g. Alzheimer's disease, Parkinson's disease, Pick's disease, Huntington's disease), infectious diseases (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or toxicoisis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism, intoxication by drugs, metal and organic compounds), tumorigenic diseases (e.g. brain tumour), traumatic diseases (e.g. chronic subarachnoidal haemorrhage, and other types of dementia, depression, hyperactive child syndrome (microencephalopathy) and disturbance of consciousness. It is also useful for prevention and treatment of diseases associated with prolactin hypo and hypersecretion respectively, including: hyperprolactinaemia, pituitary adenoma, breast cancer, infertility, impotence and autoimmune disease (hypersecretion disorders), and seminal vesicle hypoplasia,

CC osteoporosis, menopausal syndrome and renal failure (hypersecretion  
CC disorders). The 19P2 polypeptide/amide is also useful as a test  
CC reagent for study of the prolactin secretory function or as a  
CC lactogogue in mammalian farm animals.  
XX

SQ Sequence 31 AA;  
SQ

Query Match  
Best Local Similarity 96.3%; Score 105; DB 20; Length 31;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAWYXXRGIRPVGRF 20  
|||||  
DB 12 tpdinpawyasrgirpvgrf 31

Search completed: September 13, 2002, 09:18:34  
Job time: 499 sec

281 d



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:34 ; Search time 399.68 Seconds  
(without alignments)  
6.114 Million cell updates/sec

Title: US-09-446-543a-73

Perfect score: 109

Sequence: 1 TPDINPAMYXXRGIRPVGREXX 22

Scoring table: BLOSUM62

Gapop 10.0, Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: A-Geneseq\_032802:\*

1: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*

2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*

3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*

4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*

5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*

6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*

7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*

8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*

9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*

10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*

11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*

12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*

13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*

14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*

15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*

16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*

17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*

18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*

19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*

20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*

21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*

22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	105	96.3	20 18 AAW31394	Human type G prote
2	105	96.3	20 20 AAW97236	Human type G prote
3	105	96.3	20 21 AAB10365	Human type G prote
4	105	96.3	20 21 AAY49294	Human type G prote
5	105	96.3	20 22 AAG62534	Human type G prote
6	105	96.3	20 22 AAB90992	Human type G prote
7	105	96.3	21 18 AAW31395	Human type G prote
8	105	96.3	21 21 AAB10366	Human type G prote
9	105	96.3	21 22 AAG62535	Human type G prote
10	105	96.3	22 18 AAW31396	Human type G prote
11	105	96.3	22 21 AAB10367	Human type G prote

12	105	96.3	22 22 AAG62536	Human CRH releasin
13	105	96.3	31 18 AAW31391	Human type G prote
14	105	96.3	31 20 AAW97235	Human type G prote
15	105	96.3	31 20 AAB87615	Human type G prote
16	105	96.3	31 21 AAB10362	Human type G prote
17	105	96.3	31 21 AAY49291	Human type G prote
18	105	96.3	31 22 AAG62531	Human type G prote
19	105	96.3	31 22 AAB90991	Human type G prote
20	105	96.3	32 22 AAB90995	Human type G prote
21	105	96.3	32 18 AAW31392	Human type G prote
22	105	96.3	32 21 AAB10363	Human type G prote
23	105	96.3	32 22 AAG62532	Human type G prote
24	105	96.3	33 18 AAW31393	Human type G prote
25	105	96.3	33 21 AAB10364	Human type G prote
26	105	96.3	33 22 AAG62533	Human type G prote
27	105	96.3	87 18 AAW31390	Human type G prote
28	105	96.3	87 20 AAW97226	Human type G prote
29	105	96.3	87 21 AAB10361	Human type G prote
30	105	96.3	87 22 AAG62530	Human type G prote
31	104	95.4	20 18 AAW31387	Human type G prote
32	104	95.4	20 18 AAW31374	Human type G prote
33	104	95.4	20 20 AAW97232	Human type G prote
34	104	95.4	20 20 AAW97234	Human type G prote
35	104	95.4	20 20 AAW95175	Human type G prote
36	104	95.4	20 20 AAW95191	Human type G prote
37	104	95.4	20 21 AAB10350	Human type G prote
38	104	95.4	20 21 AAB10358	Human type G prote
39	104	95.4	20 21 AAB19301	Human type G prote
40	104	95.4	20 21 AAY49302	Human type G prote
41	104	95.4	20 22 AAG62519	Human type G prote
42	104	95.4	20 22 AAG62527	Human type G prote
43	104	95.4	20 22 AAB90994	Human type G prote
44	104	95.4	20 22 AAB90996	Human type G prote
45	104	95.4	20 22 AAB6954	Human type G prote

## ALIGNMENTS

RESULT 1	AAW31394	standard; Peptide: 20 AA.
ID	AAW31394	
AC	AAW31394	
XX	06-APR-1998	(first entry)
DE	Human type G protein-coupled receptor ligand fragment 4.	
KW	G protein-coupled receptor; ligand binding; pharmaceutical;	
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;	
KW	therapeutic agent.	
XX		
OS	Homo sapiens.	
PN	W09724436-A2.	
XX		
PD	10-JUL-1997.	
XX		
PF	26-DEC-1996;	96WO-JP03821.
PR	18-SEP-1996;	96JP-0246573.
PR	28-DEC-1995;	95JP-0343371.
PR	15-MAR-1996;	96JP-0059419.
PR	12-AUG-1996;	96JP-0211805.
PA	(TAKE ) TAKEDA CHEM IND LTD.	
XX		
PI	Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,	
PI	Kawamata Y, Kitada C;	
XX		
DR	WPI: 1997-363672/73.	
DR	N-PSDB; AAV02431.	

THIS PAGE BLANK (USPTO)

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 185; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
 CC sequence represented in AAW97236 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator. A central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharide. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting which are  
 CC activation of the G protein-coupled receptor protein.  
 CC  
 SQ Sequence 20 AA:

Query Match 96.3%; Score 105; DB 18; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2.6e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDIINPAMWYXXRGIRPVGRF 20  
 |||||  
 DB 1 tpdinpwasyasgrtprvgrf 20

RESULT 2  
 AAW97236  
 ID AAW97236 standard; peptide: 20 AA.

AC AAW97236;  
 DT 06-MAY-1999 (first entry)  
 DE Human type ligand polypeptide fragment.

KW Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;  
 KW lrruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.

OS Homo sapiens.  
 XX  
 XX WO9858962-A1.  
 XX  
 XX 30-DEC-1998.  
 XX  
 XX 22-JUN-1998; 98WO-JP02765.  
 XX  
 XX 23-JUN-1997; 97JP-0165437.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 XX

DR WPI: 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3; Page 166; 241pp; English.

CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenorrhoea, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC chorioncarcinoma, hydatid mole, lrruption mole, abortion, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 CC  
 SQ Sequence 20 AA:

Query Match 96.3%; Score 105; DB 20; Length 20;  
 Best Local Similarity 90.0%; Pred. No. 2.6e-11;  
 Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDIINPAMWYXXRGIRPVGRF 20  
 |||||  
 DB 1 tpdinpwasyasgrtprvgrf 20

RESULT 3  
 AAB10365  
 ID AAB10365 standard; peptide: 20 AA.

AC AAB10365;  
 DT 24-NOV-2000 (first entry)  
 DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.

KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

OS Homo sapiens.  
 XX  
 XX WO200038704-A1.  
 XX  
 XX 06-JUL-2000.  
 XX  
 XX 22-DEC-1999; 99WO-JP07199.  
 XX  
 XX 25-DEC-1998; 98JP-0369585.  
 XX  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 XX Matsumoto H, Kitada C, Hinuma S;  
 XX

WPI: 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:36 ; Search time 399.68 Seconds

(without alignments)  
3.057 Million cell updates/sec

Title: US-09-446-543a-74

Perfect score: 47

Sequence: 1 SRXHXSMEXR 11

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Lasting first 45 summaries

AGeneSeq\_032802:\*

1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	43	91.5	31	AAW31384
2	43	91.5	31	AAW97233
3	43	91.5	31	AAW87614
4	43	91.5	31	AAW95173
5	43	91.5	31	AAW95174
6	43	91.5	31	AAW95175
7	43	91.5	31	AAW95176
8	43	91.5	31	AAW95177
9	43	91.5	31	AAW95178
10	43	91.5	31	AAW95179
11	43	91.5	31	AAW95180

12	43	91.5	32	21	AAW10356	Rat oxytocin secre
13	43	91.5	32	22	AAW62525	Rat CRH releasing
14	43	91.5	33	18	AAW31386	Rat type G protein
15	43	91.5	33	21	AAW10357	Rat oxytocin secre
16	43	91.5	33	22	AAW62526	Rat CRH releasing
17	43	91.5	82	20	AAW95172	Murine pituitary-d
18	43	91.5	83	18	AAW31383	Rat type G protein
19	43	91.5	83	20	AAW97225	Rat type ligand po
20	43	91.5	83	21	AAW10354	Rat oxytocin secre
21	43	91.5	83	22	AAW62523	Rat CRH releasing
22	42	89.4	15	18	AAW31399	Synthetic ligand 1
23	42	89.4	15	20	AAW97229	N-terminal ligand
24	42	89.4	29	18	AAW31369	Bovine G protein-c
25	42	89.4	29	20	AAW95184	Bovine pituitary-d
26	42	89.4	30	21	AAW49299	19P2 ligand peptid
27	42	89.4	31	18	AAW31391	Human type G prote
28	42	89.4	31	18	AAW31371	Human type G prote
29	42	89.4	31	20	AAW97235	Human type ligand
30	42	89.4	31	20	AAW97218	Bovine pituitary-d
31	42	89.4	31	20	AAW87613	Bovine 19P2 ligand
32	42	89.4	31	20	AAW87615	Bovine pituitary-d
33	42	89.4	31	21	AAW95188	Bovine pituitary-d
34	42	89.4	31	21	AAW10347	Bovine oxytocin se
35	42	89.4	31	21	AAW10362	Human oxytocin sec
36	42	89.4	31	21	AAW49290	Human oxytocin sec
37	42	89.4	31	21	AAW49291	19P2 ligand peptid
38	42	89.4	31	21	AAW49298	19P2 ligand peptid
39	42	89.4	31	22	AAW62516	Bovine CRH releas
40	42	89.4	31	22	AAW62531	Human CRH releas
41	42	89.4	31	22	AAW90991	Prolactin releas
42	42	89.4	31	22	AAW90995	Prolactin releas
43	42	89.4	31	22	AAW73370	19P2 ligand peptid
44	42	89.4	32	18	AAW31392	Human type G prote
45	42	89.4	32	18	AAW31372	Bovine G protein-c

## ALIGNMENTS

RESULT 1	AAW31384	standard; Peptide; 31 AA.
ID	AAW31384;	
AC	AAW31384;	
XX		
DT	06-APR-1998 (first entry)	
XX		
DE	Rat type G protein-coupled receptor ligand fragment 1.	
XX		
KW	G protein-coupled receptor; ligand binding; pharmaceutical;	
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;	
KW	therapeutic agent.	
XX		
OS	Rat sp.	
XX		
PN	WO9724436-A2.	
XX		
PD	10-JUL-1997.	
XX		
PF	26-DEC-1996; 96WO-JP03821.	
XX		
PR	18-SEP-1996; 96JP-0246573.	
PR	28-DEC-1995; 95JP-0343371.	
PR	15-MAR-1996; 96JP-0059419.	
PR	12-NOV-1996; 96JP-0211803.	
XX		
PA	(TAKE ) TAKEDA CHEM IND LTD.	
PI	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;	
PI	Kawamata Y, Kitada C;	
XX		
DR	WPI; 1997-363672/33.	
DR	N-PSDB; AAW02421.	

SEQ ID NO: 74  
AC: NO: AAW31384  
Database: A-GeneSeq-032802

**THIS PAGE BLANK (USPTO)**

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 179; 258pp; English.

CC This sequence represents a peptide fragment from a novel rat type  
 CC ligand polypeptide corresponding to amino acid residues 22 to 52 of the  
 CC sequence represented in AAM97233 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trismus, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharide. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.

SO Sequence 31 AA;

Query Match 91.5%; Score 43; DB 18; Length 31;  
 Best Local Similarity 72.7%; Pred. No. 0.026;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
 ||| |||||  
 Db 1 strahqmetr 11

RESULT 2

AAM97233 standard; peptide; 31 AA.

AC AAM97233:

DT 06-MAY-1999 (first entry)

DE Rat type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;

KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;

KW menopaual syndrome; euthyroid; hypometabolism; lactation;

KW pituitary adenomatosis; brain tumour; emmenorrhoea; galactorrhoea;

KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhoea;

KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;

KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KW contraceptive; placental function; choriocarcinoma; hydralid mole;

KW abortion; uterine function; abnormal lipidmetabolism;

KW abnormal lipidmetabolism; oxytocia.

XX Rattus sp.

OS W0985962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPI: 1999-105614/09.  
 XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopaual  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3; Page 153; 241pp; English.

CC The present sequence represents a rat type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyst cacosgenesis, menopaual  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenorrhoea, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhoea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydralid mole, abortion, uterine function, abnormal  
 CC lipidmetabolism, abnormal lipidmetabolism or oxytocia.

SO Sequence 31 AA;

Query Match 91.5%; Score 43; DB 20; Length 31;  
 Best Local Similarity 72.7%; Pred. No. 0.026;  
 Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
 ||| |||||  
 Db 1 strahqmetr 11

RESULT 3

AAM87614 standard; Peptide; 31 AA.

AC AAM87614:

DT 29-MAR-1999 (first entry)

DE Rat 19P2 ligand.

XX 19P2 ligand; G protein coupled receptor; pituitary;

KW prolactin releasing peptide; rat; dementia; breast cancer;

XX therapy.

OS Rattus sp.

XX EP887417-A2.

XX 30-DEC-1998.

XX 25-JUN-1998; 98EP-0111725.

XX 27-JUN-1997; 97JP-0172118.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;

XX WPI: 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage

XX of a fusion protein, useful for preventing and treating dementia,

XX breast cancer, renal failure and autoimmune disease

XX Claim 5; Page 34; 56pp; English.

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Comugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:06:40 : Search time 30.03 Seconds  
(without alignments)  
122.059 Million cell updates/sec

Title: SEQ74-FUSED-TO-SEQ73

Perfect score: 156  
Sequence: 1 SRXHXSMEXRTPDINPAWYXXRGIRPYGRFX 33

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

A.GeneSeq\_032802:\*

1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	147	94.2	31	18	AAW31391
2	147	94.2	31	18	AAW31384
3	147	94.2	31	20	AAW97233
4	147	94.2	31	20	AAW97235
5	147	94.2	31	20	AAW87614
6	147	94.2	31	20	AAW87615
7	147	94.2	31	20	AAW5173
8	147	94.2	31	20	AAW95174
9	147	94.2	31	21	AAAB10355
10	147	94.2	31	21	AAAB10362
11	147	94.2	31	21	AAW87504

12	147	94.2	31	21	AAV49291	19P2 ligand peptid
13	147	94.2	31	21	AAV49292	19P2 ligand peptid
14	147	94.2	31	22	AAW62524	Rat CRH releasing
15	147	94.2	31	22	AAW62531	Human CRH releasing
16	147	94.2	31	22	AAW90991	Proactin releasing
17	147	94.2	31	22	AAW90993	Proactin releasing
18	147	94.2	31	22	AAW90995	Proactin releasing
19	147	94.2	32	18	AAW31392	Human type G prote
20	147	94.2	32	18	AAW31385	Rat type G prote
21	147	94.2	32	21	AAAB10356	Rat oxytocin secre
22	147	94.2	32	21	AAAB10363	Human oxytocin sec
23	147	94.2	32	22	AAW62525	Rat CRH releasing
24	147	94.2	32	22	AAW62532	Human CRH releasing
25	147	94.2	33	18	AAW31393	Human type G prote
26	147	94.2	33	18	AAW31386	Rat type G prote
27	147	94.2	33	21	AAAB10357	Rat oxytocin secre
28	147	94.2	33	21	AAAB10364	Human oxytocin sec
29	147	94.2	33	22	AAW62526	Rat CRH releasing
30	147	94.2	33	22	AAW62533	Human CRH releasing
31	147	94.2	82	20	AAW95172	Murine pituitary-d
32	147	94.2	83	18	AAW31383	Rat type G prote
33	147	94.2	83	20	AAW97225	Rat type ligand po
34	147	94.2	83	21	AAAB10354	Rat oxytocin secre
35	147	94.2	83	22	AAW62523	Rat CRH releasing
36	147	94.2	87	18	AAW31390	Human type G prote
37	147	94.2	87	20	AAW97226	Human type ligand
38	147	94.2	87	21	AAAB10361	Human oxytocin sec
39	147	94.2	87	22	AAW62530	Human CRH releasing
40	146	93.6	31	18	AAW31371	Bovine G prote
41	146	93.6	31	20	AAW97218	Bovine pituitary-d
42	146	93.6	31	20	AAW87613	Bovine 19P2 ligand
43	146	93.6	31	20	AAW95188	Bovine pituitary-d
44	146	93.6	31	21	AAAB10347	Bovine oxytocin se
45	146	93.6	31	21	AAV49290	19P2 ligand peptid

#### ALIGNMENTS

RESULT	ID	AAW31391 standard; Peptide: 31 AA.
AC	AAW31391:	
DT	06-APR-1998 (first entry)	
DE	Human type G protein-coupled receptor ligand fragment 1.	
XX		
KW	G protein-coupled receptor; ligand binding; pharmaceutical;	
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;	
XX	therapeutic agent.	
OS	Homo sapiens.	
XX		
XX	WO9724436-A2.	
PD	10-JUL-1997.	
XX		
PF	26-DEC-1996; 96WO-JP03821.	
XX		
PR	18-SEP-1996; 96JP-0246573.	
PR	28-DEC-1995; 95JP-0343371.	
PR	15-MAR-1996; 96JP-0059419.	
PR	12-AUG-1996; 96JP-0211805.	
XX		
PA	(TAKE ) TAKEDA CHEM IND LTD.	
XX		
PI	Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;	
PI	Kawamata Y, Kitada C;	
XX		
DR	WPI, 1997-363672/33.	
DR	N-PSDB; AAV02428.	

**THIS PAGE BLANK (USPTO)**

**SQ Sequence 31 AA;**

QY 1 SRXHXHSMEXRTPDINPAWYXXRGIRPVGRF 31

Db 1 srthrhsmelrtptdlnpawyasrglrpvgrf 31

```

RESULT 2
AAW31384
ID AAW31384 standard; peptide; 31 AA.
vv

```

AC AAW31384

DT 06-APR-1998 (first entry)

Rat type G protein-coupled receptor ligand fragment 1. ....

KM G protein-coupled receptor; ligand binding; pharmaceutical;  
 KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KM therapeutic agent.

Rat sp.

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 15-MAR-1996; 96JP-0059419.

XX  
XX

[illegible]

PI Kawamata Y, Kitada C;  
yy

DR WPI; 1997-363672/33.

XX

PT Ligand-peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2, page 179; 258pp; English.

CC This sequence represents a peptide

ligand polypeptide corresponding to amino acid residues 22 to 52 of the sequence represented in AAW3133 and is used in an assay to monitor ligand binding to the G protein-coupled receptor protein. Pharmaceutical compositions containing this ligand may be used as a pituitary function modulator, a central nervous system modulator or a pancreatic function modulator. This ligand could have specific applications as a prophylactic or therapeutic agent for dementia, depression, hyperkinetic syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia, trauma, growth hormone secretory disease, hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia, hyperglycaemia, hyperproliferative diseases, diabetes, cancer, pancreaticitis, renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis, acute myocardial infarction, infertility, spinocerebellar degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis and/or oligosaccharia. Assays can also be developed to screen compounds which are capable of altering the binding activity of the ligand affecting activation of the G protein-coupled receptor protein.

**SQ Sequence 31 AA;**

Query Match	94.28;	Score 147;	DB 18;	Length 31;
Best Local Similarity	83.98;	Pred. No. 2.3e-17;		
Matches 26; Conservative	0;	Mismatches 5;	Indels 0;	Gaps 0

QY 1 SRXHXHSMEXRTPDINPAWYXXRCIRPVGRF 31

Db 1 strahqhsmetrtplnpawytgrgjrpgvf 31

RESULT	3
AAW97233	
ID	AAW97233 standard; peptide; 31 AA
XX	

AC 4449 / 6000  
XX

00-MAR-1959 (LIST ONLY)  
XX

DE Rac type ligand polypeptide fragments  
XX

KW G protein-coupled receptor; GPCR; hypoovarianism; donecyst cacogenesls;

KW pituitary adenomatosis; brain tumour; autoimmune disease;

KW proactinoma; interitilly; impotence; amenorrhea; galactorrhea;  
acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW

KW forbes-aldridge syndrome; lymphoma; sneehan syndrome; dyszoospermia; contraceptive; placental function; chorlocarcinoma; hydatid mole;

KW abnormal lipidmetabolism; oxytocin.  
KW more; abolition; unclinically rectus; abnormal saccharidometabolism;  
KW

AA Rattus sp.  
OS

PN WO9858962-A1

AA 30-DEC-1998  
PD

AA 22-JUN-1998; 98WO-JP02765.  
PF

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fuji R, Hinuma S, Kawamata Y, Matsumoto H;

WPI; 1999-105614/09.

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using SW model

Run on: September 13, 2002, 09:10:15 ; Search time 399.68 Seconds  
(without alignments)  
8.615 Million cell updates/sec

Title: US-09-446-543a-61  
Perfect score: 171  
Sequence: 1 SRTNRHSMKIRTPDINPAMVSRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: A\_Geneseq\_032802.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	171	100.0	31	AAW31391	Human type G prote
2	171	100.0	31	AAW97235	Human type 1 ligand
3	171	100.0	31	AAW87615	Human 19P2 ligand
4	171	100.0	31	AAW10362	Human oxytocin sec
5	171	100.0	31	AAW49291	19P2 ligand peptid
6	171	100.0	31	AAW62531	Human CRH releasin
7	171	100.0	31	AAW90991	Proactin releasin
8	171	100.0	32	AAW31392	Human type G prote
9	171	100.0	32	AAW10363	Human oxytocin sec
10	171	100.0	32	AAW62532	Human CRH releasin
11	171	100.0	33	AAW31393	Human type G prote

12	171	100.0	33	21	AAW10364	Human oxytocin sec
13	171	100.0	33	22	AAW62533	Human CRH releasin
14	171	100.0	87	18	AAW31390	Human type G prote
15	171	100.0	87	20	AAW97236	Human type 1 ligand
16	171	100.0	87	21	AAW10361	Human oxytocin sec
17	171	100.0	87	22	AAW62530	Human CRH releasin
18	165	96.5	30	21	AAW49299	19P2 ligand peptid
19	162	94.7	31	22	AAW90995	Proactin releasin
20	158	92.4	31	18	AAW31371	Bovine G protein-c
21	158	92.4	31	20	AAW97218	Bovine pituitary-d
22	158	92.4	31	20	AAW87613	Bovine 19P2 ligand
23	158	92.4	31	21	AAW95188	Bovine pituitary-d
24	158	92.4	31	21	AAW10347	Bovine oxytocin se
25	158	92.4	31	21	AAW49290	19P2 ligand peptid
26	158	92.4	31	21	AAW49298	19P2 ligand peptid
27	158	92.4	31	22	AAW62516	Bovine CRH releasi
28	158	92.4	32	18	AAW31372	Bovine CRH releasi
29	158	92.4	32	20	AAW95189	Bovine G protein-c
30	158	92.4	32	21	AAW10348	Bovine pituitary-d
31	158	92.4	32	22	AAW62517	Bovine oxytocin se
32	158	92.4	33	18	AAW31373	Bovine CRH releasi
33	158	92.4	33	20	AAW95190	Bovine G protein-c
34	158	92.4	33	21	AAW10349	Bovine pituitary-d
35	158	92.4	33	21	AAW49297	Bovine oxytocin se
36	158	92.4	33	22	AAW62518	19P2 ligand peptid
37	158	92.4	98	18	AAW31382	Bovine CRH releasi
38	158	92.4	98	18	AAW31368	Bovine G protein-c
39	158	92.4	98	20	AAW97224	Bovine genome-driv
40	158	92.4	98	20	AAW97217	Bovine pituitary-d
41	158	92.4	98	20	AAW95187	Bovine genome-driv
42	158	92.4	98	21	AAW10346	Bovine oxytocin se
43	158	92.4	98	21	AAW10353	Bovine oxytocin se
44	158	92.4	98	22	AAW62515	Bovine CRH releasi
45	158	92.4	98	22	AAW62522	Bovine CRH releasi

## ALIGNMENTS

RESULT	1	ALIGNMENTS
AAW31391	standard; Peptide; 31 AA.	
AAW31391		
AAW31391	(first entry)	
06-APR-1998		
Human type G protein-coupled receptor ligand fragment 1.		
G protein-coupled receptor; ligand binding; pharmacological;		
modulator; pituitary; central nervous system; pancreas; prophylactic;		
therapeutic agent.		
Homo sapiens.		
WO9724436-A2.		
10-JUL-1997.		
26-DEC-1996;	96MO-JF03821.	
18-SEP-1996;	96JP-0246573.	
28-DEC-1995;	95JP-0343371.	
15-MAR-1996;	96JP-0059419.	
12-AUG-1996;	96JP-0211805.	
(TAKE ) TAKEDA CHEM IND LTD.		
Fujita R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;		
Kawamata Y, Kltada C;		
WPI; 1997-363672/33.		
N-PDB; AAW02428.		

**THIS PAGE BLANK (USPTO)**

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 PS Claim 2; Page 184; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
 CC sequence represented in AAM97235 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidemia, hypercholesterolemia, hyperglycemia,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharide. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.

SQ Sequence 31 AA;

Query Match 100.0%; Score 171; DB 18; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDINPAWASRGIRPVGRF 31  
 DB 1 SRTHRSMETRPDINPAWASRGIRPVGRF 31

RESULT 2

AAM97235  
 ID AAM97235 standard; peptide: 31 AA.

AC AAM97235;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

XX Rat type ligand; modulation: prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypocoarctanism; gonocyst cacosgenesis;  
 KW menopause; endometriosis; brain tumour; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; amenorrhoea; galactorrhoea;  
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhoea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
 KW interruption; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.

OS Homo sapiens.

XX WO9858962-A1.

PN 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPT; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopause  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy

PS Claim 3; Page 159; 241pp; English.

CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypocoarctanism, gonocyst cacosgenesis, menopause  
 CC syndrome, endometriosis or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, amenorrhoea, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhoea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dysospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC chorocarcinoma, hydralid mole, abortion, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

SQ Sequence 31 AA;

Query Match 100.0%; Score 171; DB 20; Length 31;  
 Best Local Similarity 100.0%; Pred. No. 4.1e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDINPAWASRGIRPVGRF 31  
 DB 1 SRTHRSMETRPDINPAWASRGIRPVGRF 31

RESULT 3

AAM87615  
 ID AAM87615 standard; peptide: 31 AA.

AC AAM87615;

DT 29-MAR-1999 (first entry)

DE Human 19P2 ligand.

XX 19P2 ligand; G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; human; dementia; breast cancer;  
 KW therapy.

OS Homo sapiens.

XX EP887417-A2.

XX 30-DEC-1998.

XX 25-JUN-1998; 98EP-0111725.

XX 27-JUN-1997; 97JP-0172118.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;

XX WPT; 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
 PT of a fusion protein, useful for preventing and treating dementia,  
 PT breast cancer, renal failure and autoimmune disease  
 XX Claim 5; Page 35; 56pp; English.

**THIS PAGE BLANK (USPTO)**



check 3 & 4 date

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:57 ; Search time 136.62 Seconds  
(without alignments)  
3.933 Million cell updates/sec

Title: US-09-446-543a-73  
Perfect score: 109  
Sequence: 1 TPDIHPAMYXXRGIRPVGREXX 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5a\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5b\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6a\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6b\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/pctus\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	105	96.3	20	3	US-09-105-678A-46
2	105	96.3	20	4	US-08-776-971-64
3	105	96.3	20	4	US-09-421-208-46
4	105	96.3	21	3	US-09-105-678A-47
5	105	96.3	21	4	US-08-776-971-65
6	105	96.3	21	4	US-09-421-208-47
7	105	96.3	22	3	US-09-105-678A-48
8	105	96.3	22	4	US-08-776-971-66
9	105	96.3	22	4	US-09-421-208-48
10	105	96.3	31	3	US-09-105-678A-9
11	105	96.3	31	3	US-09-105-678A-43
12	105	96.3	31	4	US-08-776-971-61
13	105	96.3	31	4	US-09-421-208-9
14	105	96.3	31	4	US-09-105-678A-44
15	105	96.3	32	3	US-09-105-678A-45
16	105	96.3	32	4	US-08-776-971-62
17	105	96.3	32	4	US-09-421-208-44
18	105	96.3	33	3	US-09-105-678A-45
19	105	96.3	33	4	US-08-776-971-63
20	105	96.3	33	4	US-09-421-208-45
21	105	96.3	87	4	US-08-776-971-59
22	105	96.3	87	4	US-08-776-971-135
23	105	96.3	87	4	US-08-776-971-138
24	104	95.4	20	3	US-09-105-678A-34
25	104	95.4	20	3	US-09-105-678A-40
26	104	95.4	20	4	US-08-776-971-8
27	104	95.4	20	4	US-08-776-971-50

28	104	95.4	20	4	US-08-776-971-98	Sequence 98, Appl
29	104	95.4	20	4	US-09-421-208-34	Sequence 34, Appl
30	104	95.4	20	4	US-09-421-208-40	Sequence 40, Appl
31	104	95.4	21	3	US-09-105-678A-35	Sequence 35, Appl
32	104	95.4	21	3	US-09-105-678A-41	Sequence 41, Appl
33	104	95.4	21	4	US-08-776-971-9	Sequence 9, Appl1
34	104	95.4	21	4	US-08-776-971-51	Sequence 51, Appl
35	104	95.4	21	4	US-09-421-208-35	Sequence 35, Appl
36	104	95.4	21	4	US-09-421-208-41	Sequence 41, Appl
37	104	95.4	22	3	US-09-105-678A-36	Sequence 36, Appl
38	104	95.4	22	3	US-09-105-678A-42	Sequence 42, Appl
39	104	95.4	22	4	US-08-776-971-10	Sequence 10, Appl
40	104	95.4	22	4	US-08-776-971-52	Sequence 52, Appl
41	104	95.4	22	4	US-09-421-208-36	Sequence 36, Appl
42	104	95.4	22	4	US-09-421-208-42	Sequence 42, Appl
43	104	95.4	31	3	US-09-105-678A-7	Sequence 7, Appl1
44	104	95.4	31	3	US-09-105-678A-8	Sequence 8, Appl1
45	104	95.4	31	3	US-09-105-678A-31	Sequence 31, Appl

ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Walter Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 96.3% Score 105; DB 3; Length 20;  
Best Local Similarity 90.0% Pred. No. 1.8e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXRGIRPVGRF 20  
|||||  
Db 1 TPDINPAMYASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64

; Sequence 64, Application US/0876971B  
; Patent No. 6228984

## GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

Hinata, Yugo

Kawamata, Yuji

Hosoya, Masaki

Fuji, Ryo

Fukusumi, Shoji

Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM compatible

OPERATING SYSTEM: DOS

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: &lt;Unknown&gt;

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein

FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 64:

US-08-776-971-64

Query Match 96.3%; Score 105; DB 4; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.8e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXRGIRPVGRF 20  
|||||  
Db 1 TPDINPAMYASRGIRPVGRF 20

## RESULT 3

US-09-421-208-46

; Sequence 46, Application US/09421208  
; Patent No. 6258561

## GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

Moriya, Takeo

Tanaka, Yoko

Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,208

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678

FILING DATE: 26-JUN-1998

APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:

LENGTH: 20 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-421-208-46

Query Match 96.3%; Score 105; DB 4; Length 20;  
Best Local Similarity 90.0%; Pred. No. 1.8e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXRGIRPVGRF 20  
|||||  
Db 1 TPDINPAMYASRGIRPVGRF 20

## RESULT 4

US-09-105-678A-47

; Sequence 47, Application US/09105678A  
; Patent No. 6103882

## GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

Moriya, Takeo

Tanaka, Yoko

Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS &amp; CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105, 678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 96.3%; Score 105; DB 3; Length 21;  
Best Local Similarity 90.0%; Pred. No. 1.9e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 1 TPDINPAMYASRGIRPVGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habela, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 96.3%; Score 105; DB 4; Length 21;  
Best Local Similarity 90.0%; Pred. No. 1.9e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 1 TPDINPAMYASRGIRPVGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421, 208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105, 678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 96.3%; Score 105; DB 4; Length 21;  
Best Local Similarity 90.0%; Pred. No. 1,9e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDINPAWYXRGIRPVGRF 20  
||||||| |||||||  
DB 1 TPDINPAWYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 96.3%; Score 105; DB 3; Length 22;  
Best Local Similarity 90.0%; Pred. No. 2e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDINPAWYXRGIRPVGRF 20  
||||||| |||||||  
DB 1 TPDINPAWYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/21805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:

US-08-776-971-66

Query Match 96.3%; Score 105; DB 4; Length 22;  
Best Local Similarity 90.0%; Pred. No. 2e-11;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 TPDINPAWYXRGIRPVGRF 20  
||||||| |||||||  
DB 1 TPDINPAWYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

```
; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; US-09-421-208-48
;
; Query Match          96.3%; Score 105; DB 4; Length 22;
; Best Local Similarity 90.0%; Pred. No. 2e-11;
; Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TPDINPAMVXXRGIRPYGRF 20
    ||||||| |||||||
Db 1 TPDINPAMVASRGIRPYGRF 20
```

```
RESULT 10
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
;
; US-09-105-678A-9
```

```
Query Match          96.3%; Score 105; DB 3; Length 31;
Best Local Similarity 90.0%; Pred. No. 3e-11;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 TPDINPAMVXXRGIRPYGRF 20
    ||||||| |||||||
Db 12 TPDINPAMVASRGIRPYGRF 31
```

```
RESULT 11
US-09-105-678A-43
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
```

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-43

Query Match  
Best Local Similarity 96.3%; Score 105; DB 3; Length 31;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 12 TPDINPAMYASRGIRPVGRF 31

RESULT 12  
US-08-776-971-61  
Sequence 61, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match  
Best Local Similarity 96.3%; Score 105; DB 4; Length 31;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 12 TPDINPAMYASRGIRPVGRF 31

RESULT 13  
US-09-421-208-9  
Sequence 9, Application US/09421208  
Patent No. 6258361  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
Applicant: Tanaka, Yoko  
Applicant: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match  
Best Local Similarity 96.3%; Score 105; DB 4; Length 31;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 12 TPDINPAMYASRGIRPVGRF 31

RESULT 14

```

US-09-421-208-43
: Sequence 43, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: NUMBER OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-09-421-208-43

Query Match          96.3%; Score 105; DB 4; Length 31;
Best Local Similarity 90.0%; Pred. No. 3e-11;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY      1  TPDINPAMYXXRGIRPVGRF 20
        ||||||| |||||||
Db      12  TPDINPAMYASRGIRPVGRF 31

RESULT 15
US-09-105-678A-44
: Sequence 44, Application US/09105678A
: Patent No. 6103882
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA

```

```

? ZIP: 02109
? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patentin Release #1.0, Version #1.30
? CURRENT APPLICATION DATA:
? APPLICATION NUMBER: US/09/105,678A
? FILING DATE: 26-JUN-1998
? PRIOR APPLICATION DATA:
? APPLICATION NUMBER: JP 172118/1997
? FILING DATE: 27-JUN-1997
? ATTORNEY/AGENT INFORMATION:
? NAME: Conlin, David G.
? REGISTRATION NUMBER: 27,026
? REFERENCE/DOCKET NUMBER: 48466-342
? TELECOMMUNICATION INFORMATION:
? TELEPHONE: 617-523-3400
? TELEFAX: 617-523-6440
? INFORMATION FOR SEQ ID NO: 44:
? SEQUENCE CHARACTERISTICS:
? LENGTH: 32 amino acids
? TYPE: amino acid
? STRANDEDNESS:
? TOPOLOGY: linear
? MOLECULE TYPE: peptide
? US-09-105-678A-44
?
Query Match          96.3%, Score 105; DB 3; Length 32;
Best Local Similarity 90.0%; Pred. No. 3.le-11;
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
OY      1 TPDINPAMYXXRGIRPVGRF 20
        |||||||  |||||||
Db       12 TPDINPAWYASRGIRPVGRF 31

```

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:23:57 ; Search time 172.41 Seconds  
(without alignments)  
12.261 Million cell updates/sec

Title: US-09-446-543A-73

Sequence: 1 TPDINPAMYXXRGIRPVGRFX 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	104	95.4	83	2 JC7607	prolactin-releasing
2	49.5	45.4	664	2 F83376	conserved hypothet
3	49	45.0	790	2 T47959	hypothetical prote
4	46	42.2	333	2 H82852	hydroxybenzoate oc
5	45	41.3	767	2 T21969	hypothetical prote
6	44	40.4	250	2 G83400	hypothetical prote
7	44	40.4	284	2 F71015	hypothetical prote
8	43.5	39.9	1501	2 T45623	hypothetical prote
9	43	39.4	220	2 C83292	probable glutathio
10	43	39.4	232	2 G75608	hypothetical prote
11	43	39.4	309	2 T32576	hypothetical prote
12	43	39.4	376	2 A48197	opsin, ocellar - A
13	43	39.4	376	2 B48197	opsin, lateral eye
14	43	39.4	476	2 G64720	probable amino aci
15	43	39.4	476	2 G90629	probable inner mem
16	43	39.4	476	2 G85480	inner membrane tra
17	43	39.4	719	2 S61046	ARPI protein - yea
18	43	39.4	1296	2 T16859	hypothetical prote
19	42.5	39.0	443	2 T21499	hypothetical prote
20	42	38.5	428	2 F81694	pyruvate dehydroge
21	42	38.5	21	2 H87660	peptidoglycan-bind
22	41.5	38.1	345	2 DB4012	N-acetylglutamate
23	41	37.6	128	2 S76855	hypothetical prote
24	41	37.6	226	2 A87664	hypothetical prote
25	41	37.6	284	2 A75117	hypothetical prote
26	41	37.6	338	2 T20100	hypothetical prote
27	41	37.6	342	2 B64395	malic acid transpo
28	41	37.6	343	2 T46534	probable PMN-depen
29	41	37.6	347	2 H64371	malic acid transpo

30	41	37.6	419	2 AH3166	hypothetical prote
31	41	37.6	476	2 AG0502	probable amino-aci
32	41	37.6	545	2 AB7448	conserved hypothet
33	41	37.6	986	1 OXURGA	spectact receptor p
34	41	37.6	2155	2 T30197	alpha tectorin - m
35	40.5	37.2	779	2 T49717	related to BCS1 pr
36	40.5	37.2	1540	2 T45619	hypothetical prote
37	40	36.7	184	2 T35841	probable membrane
38	40	36.7	324	2 T35901	probable arac faml
39	40	36.7	329	2 H70744	hypothetical prote
40	40	36.7	341	2 T35426	probable oxidoredu
41	40	36.7	359	2 T40084	PWMP domain protei
42	40	36.7	390	2 G82844	cysteine synthase
43	40	36.7	430	1 B69009	conserved hypothet
44	40	36.7	462	2 T00708	violaxanthin de-ep
45	40	36.7	468	2 C83160	nitrite extrusion

#### ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R.Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YMM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 95.4%; Score 104; DB 2; Length 83;  
Best Local Similarity 90.0%; Pred. No. 5.1e-10;  
Matches 18; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 TPDINPAMYXXRGIRPVGRF 20  
DB 33 TPDINPAMYXXRGIRPVGRF 52

RESULT 2  
F83376  
conserved hypothetical protein PA2151 [Imported] - Pseudomonas aeruginosa (strain PAO  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83376  
R.Stover, C.K.; Pham, X.Q.; Ewlin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lardig, K.; L.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa  
A:Reference number: AB2950; MUID:20437337  
A:Accession: F83376  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STO>  
A:Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AA05539.1; GSPDB:GN  
C:Genetics:  
A:Gene: PA2151





Query Match 39.4%; Score 43; DB 2; Length 309;  
Best Local Similarity 44.4%; Pred. No. 21;  
Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

OY 2 PDINPAWYXXRIGR 19  
||:|:|:|:|:|:  
DB 131 PDSPWYLPKS--PIGR 146

RESULT 12  
A48197  
opsin, ocellar - Atlantic horseshoe crab  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
C:Accession: A48197  
R:Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
A:Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, Limulus polyphemus  
A:Reference number: A48197; MUID:93317641  
A:Accession: A48197  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <SMT>  
A:Cross-references: EMBL:U03792; NID:g156644; PIDN:AAA28274.1; PID:g156645  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: Chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; retin  
F:318/Binding site: retinal (Lys) (covalent) #status predicted

Query Match 39.4%; Score 43; DB 2; Length 376;  
Best Local Similarity 46.2%; Pred. No. 26;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 PDINPAWYXXRIG 14  
|:|:|:|:|:|:  
DB 40 PMNPLMYISILGV 52

RESULT 13  
B48197  
opsin, lateral eye - Atlantic horseshoe crab  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
C:Accession: B48197  
R:Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
A:Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, Limulus polyphemus  
A:Reference number: A48197; MUID:93317641  
A:Accession: B48197  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <SMT>  
A:Cross-references: EMBL:U03791; NID:g156642; PIDN:AAA28273.1; PID:g156643  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; retin  
F:318/Binding site: retinal (Lys) (covalent) #status predicted

Query Match 39.4%; Score 43; DB 2; Length 376;  
Best Local Similarity 46.2%; Pred. No. 26;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 PDINPAWYXXRIG 14  
|:|:|:|:|:|:  
DB 40 PMNPLMYISILGV 52

RESULT 14  
G64720  
probable amino acid transport protein yaad, sodium-dependent - Escherichia coli  
C:Species: Escherichia coli

C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
C:Accession: G64720  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G64720  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <BLAT>  
A:Cross-references: GB:AE000111; GB:U00096; NID:g1786181; PIDN:ACC73118.1; PID:g17861  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yaad  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein  
C:Keywords: amino acid transport; transmembrane protein  
F:10-26/Domain: transmembrane #status predicted <TM1>  
F:91-107/Domain: transmembrane #status predicted <TM2>  
F:142-158/Domain: transmembrane #status predicted <TM3>  
F:178-194/Domain: transmembrane #status predicted <TM4>  
F:208-224/Domain: transmembrane #status predicted <TM5>  
F:303-319/Domain: transmembrane #status predicted <TM6>  
F:349-365/Domain: transmembrane #status predicted <TM7>  
F:391-407/Domain: transmembrane #status predicted <TM8>  
F:414-430/Domain: transmembrane #status predicted <TM9>

Query Match 39.4%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 32;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 3 DIN-----PAWYXXRIG 14  
|:|:|:|:|:|:  
DB 120 DVNGQFRCGPAPWYMARGL 137

RESULT 15  
G90629  
probable inner membrane transport protein Ecs0007 [imported] - Escherichia coli (stra  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: G90629  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C  
gasawara, N.; Yasunaga, T.; Kuhtara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90629  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA833430.1; PID:g13359463; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: Ecs0007  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 39.4%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 32;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 3 DIN-----PAWYXXRIG 14  
|:|:|:|:|:|:  
DB 120 DVNGQFRCGPAPWYMARGL 137

Search completed: September 13, 2002, 09:23:58  
Job time: 773 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:26 ; Search time 311.85 Seconds  
(without alignments)  
6.102 Million cell updates/sec

Title: US-09-446-543a-74  
Perfect score: 47  
Sequence: 1 SRXHXHSMEXR 11

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPRMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriopl:\*  
17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	35	74.5	450	4 095367	095367 homo sapien
2	35	74.5	451	4 012804	012804 homo sapien
3	33	70.2	340	2 09RH19	09RH19 zymomonas m
4	33	70.2	4133	13 09DE12	09DE12 gallus galli
5	32	68.1	196	17 09HNPA	09HNPA halobacteri
6	32	68.1	277	11 09DBA2	09DBA2 mus musculu
7	32	68.1	286	11 09DAE5	09DAE5 mus musculu
8	32	68.1	383	16 084156	084156 chlamydia t
9	32	68.1	450	11 09DA19	09DA19 mus musculu
10	32	68.1	648	16 09XOW5	09XOW5 thermocoga
11	31	66.0	165	2 09JN22	09JN22 agrobacteri
12	31	66.0	756	5 09UL16	09UL16 drosophila
13	31	66.0	800	5 09V731	09V731 drosophila
14	31	66.0	905	5 09W178	09W178 drosophila
15	30	63.8	147	5 09B168	09B168 blackburnia
16	30	63.8	274	10 0949G1	0949G1 oryza sativ

17	30	63.8	330	16 09K7F3	09K7F3 bacillus ha
18	30	63.8	330	16 09A0B5	09A0B5 streptococ
19	30	63.8	332	16 005268	005268 bacillus su
20	30	63.8	393	5 023359	023359 caenorhabdi
21	30	63.8	394	10 09A5A8	09A5A8 oryza sativ
22	30	63.8	462	11 09C107	09C107 mus musculu
23	30	63.8	509	2 09XA23	09XA23 streptomyce
24	30	63.8	528	3 09C260	09C260 neurospora
25	30	63.8	599	16 09A0U0	09A0U0 oryza sativ
26	30	63.8	707	10 09AUS0	09AUS0 oryza sativ
27	30	63.8	1889	10 094H07	094H07 oryza sativ
28	29	61.7	96	8 094Y20	094Y20 pyriatella 1
29	29	61.7	122	16 09K0W8	09K0W8 neisseria m
30	29	61.7	122	16 09J710	09J710 neisseria m
31	29	61.7	160	11 091V25	091V25 mus musculu
32	29	61.7	194	10 041320	041320 secale cere
33	29	61.7	220	12 0903U4	0903U4 cymbidium m
34	29	61.7	233	3 09P826	09P826 candida alb
35	29	61.7	336	10 09AV29	09AV29 oryza sativ
36	29	61.7	438	5 020938	020938 caenorhabdi
37	29	61.7	447	2 09F7M8	09F7M8 uncultured
38	29	61.7	445	5 0960F3	0960F3 drosophila
39	29	61.7	506	11 062157	062157 mus musculu
40	29	61.7	511	5 09W2C9	09W2C9 drosophila
41	29	61.7	511	5 09N9Y8	09N9Y8 drosophila
42	29	61.7	513	11 0991K1	0991K1 mus musculu
43	29	61.7	541	11 09JK89	09JK89 mus musculu
44	29	61.7	569	16 09RU39	09RU39 delnecoccus
45	29	61.7	774	4 09BW70	09BW70 homo sapien

## ALIGNMENTS

RESULT 1  
ID 095367 PRELIMINARY; PRT; 450 AA.  
AC 095367;  
DT 01-MAY-1999 (TREMBlrel. 10, Created)  
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CBFI INTERACTING COREPRESSOR CIR.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99093483; PubMed=9874765;  
RA Hsieh J.-Y., Zhou S., Chen L., Young D.B., Hayward S.D.;  
RT "CIR, a corepressor linking the DNA binding factor CBFI to the histone  
RT deacetylase complex.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:23-28(1999).  
DR EMBL, AF098297; AAD05243.1;-  
SQ SEQUENCE 450 AA; 52328 MW; 96AABF69CB73357 CRC64;

Query Match 74.5%; Score 35; DB 4; Length 450;  
Best Local Similarity 63.6%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
QY 1 SRXHXHSMEXR 11  
Db 359 SRTXKHSPEKR 369  
RESULT 2  
ID 012804 PRELIMINARY; PRT; 451 AA.  
AC 012804;  
DT 01-NOV-1996 (TREMBlrel. 01, Created)  
DT 01-NOV-1996 (TREMBlrel. 01, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)

DE RECEPTIN.  
GN RECEPTIN.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Chai K.X., Li L., Chao J., Chao L.;  
RT "Receptin: a novel human liver cDNA encoding a serpin-like molecule.";  
RL Submitted (NOV-1993) to the EMBL/Genbank/DBJ databases.  
DR EMBL: U03644; AAA17853.1; -;  
DR InterPro: IPR000345; Cytochrome\_bnd.  
DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
SQ SEQUENCE 451 AA; 52942 MW; 157F04804DA0817E CRC64;

Query Match 74.5%; Score 35; DB 4; Length 451;  
Best Local Similarity 63.6%; Pred. No. 10;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
DB 360 SRXHXSMEXR 370

RESULT 3  
ID 09RH19 PRELIMINARY; PRT; 340 AA.  
AC 09RH19;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE CYTOCHROME OXIDASE D SUBUNIT B.  
GN CYDB.  
OS Zymomonas mobilis.  
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
OC Zymomonas.  
OX NCBI\_TaxID=542;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ZM4;  
RA Lee J.S., Kang H.S.;  
RT "Sequence analysis of 6563 cosmid clone of Zymomonas mobilis ZM4  
RT containing rRNA operon."  
RL Submitted (AUG-1998) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AF088897; AAF18273.1; -;  
DR InterPro: IPR003317; Cyto\_ox\_2.  
DR Pfam: PF02322; Cyto\_ox\_2; 1.  
SQ SEQUENCE 340 AA; 37206 MW; 3D7038A3F72EC316 CRC64;

Query Match 70.2%; Score 33; DB 2; Length 340;  
Best Local Similarity 60.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 RXXHSMEXR 11  
DB 107 RVHANSKEAR 116

RESULT 4  
ID 09DE12 PRELIMINARY; PRT; 4133 AA.  
AC 09DE12;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE DNA-DEPENDENT PROTEIN KINASE CATALYTIC SUBUNIT.  
GN DNA-PKCS.  
OS Gallus gallus (Chicken).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
OC Gallus.  
OX NCBI\_TaxID=9031;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LYMPHOCYTE;  
RA MEDLINE=20456723; PubMed=11003390;  
RA Fujimori A., Araki R., Fukumura R., Ohnata T., Takahashi H.,  
RA Kawahara A., Tatsumi K., Abe M.;  
RT "Identification of four highly conserved regions in DNA-PKcs.";  
RL Immunogenetics 51:965-973(2000).  
DR EMBL: AB016240; BAA36956.1; -;  
DR InterPro: IPR003151; FAT.  
DR InterPro: IPR003152; FATC.  
DR InterPro: IPR004043; P13\_P14\_kinase.  
DR Pfam: PF02259; FAT; 1.  
DR Pfam: PF02260; FATC; 1.  
DR Pfam: PF00454; P13\_P14\_kinase; 1.  
DR SMART; SM00146; P13K; 1.  
DR PROSITE; PS00915; P13\_4\_KINASE\_1; UNKNOWN\_1.  
DR PROSITE; PS50290; P13\_4\_KINASE\_3; 1.  
KW kinase.  
SQ SEQUENCE 4133 AA; 472522 MW; 0A15C6131466A068 CRC64;

Query Match 70.2%; Score 33; DB 13; Length 4133;  
Best Local Similarity 75.0%; Pred. No. 2,7e+02;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 RXXHSMEXR 9  
DB 1766 RDRHSME 1773

RESULT 5  
ID 09HNP4 PRELIMINARY; PRT; 196 AA.  
AC 09HNP4;  
DT 01-MAR-2001 (TREMBLrel. 16, Created)  
DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TREMBLrel. 16, Last annotation update)  
DE VNG2006C.  
GN VNG2006C.  
OS Halobacterium sp. (strain NRC-1).  
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;  
OC Halobacterium.  
OX NCBI\_TaxID=64091;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=20504483; PubMed=11016950;  
RA Ng W.V., Kennedy S.P., Mahairas G.G., Bergquist B., Pan M.,  
RA Shukla H.D., Lasly S.R., Baliga N.S., Thorsson V., Shrogha J.,  
RA Swartzell S., Weir D., Hall J., Dahl T.A., Welt R., Goo Y.A.,  
RA Leitbauer B., Keller K., Cruz R., Danson M.J., Hough D.W.,  
RA Madocks D.G., Jdionski P.E., Krebs M.P., Angeline C.M., Dale H.,  
RA Isenbacher T.A., Peck R.F., Pohlschroder M., Spudlich J.L., Jung K.-H.,  
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,  
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;  
RT "Genome sequence of Halobacterium species NRC-1.";  
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).  
DR EMBL: AE005094; AAG20176.1; -;  
DR Complete proteome.  
SQ SEQUENCE 196 AA; 21450 MW; 2F97FA24EE7C9E73 CRC64;

Query Match 68.1%; Score 32; DB 17; Length 196;  
Best Local Similarity 54.5%; Pred. No. 19;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
DB 112 SRHNSLEDR 122

```

RESULT 6
ID Q9DBA2 PRELIMINARY: PRT: 277 AA.
AC Q9DBA2:
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 1700023B02RIK PROTEIN.
GN 1700023B02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE.
RX MEDLINE=21085660; PubMed=11217851;
RA Kawal J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kodota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK008242; BAB25553.1;
DR MGD: MGI:1914185; 1700023B02RIK.
SQ
SEQUENCE 277 AA; 31797 MW; 4BEAD95516FBN32 CRC64;

Query Match 68.1%; Score 32; DB 11; Length 277;
Best Local Similarity 54.5%; Pred. No. 27;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 200 SRSRHSPEKR 210

RESULT 7
ID Q9DAE5 PRELIMINARY: PRT: 286 AA.
AC Q9DAE5:
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DE 1700023B02RIK PROTEIN.
GN 1700023B02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
ON NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawal J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaoka I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kodota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK005900; BAB24307.1;
DR MGD: MGI:1914185; 1700023B02RIK.
SQ
SEQUENCE 286 AA; 32818 MW; 068003C5E894827B CRC64;

Query Match 68.1%; Score 32; DB 11; Length 286;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 209 SRSRHSPEKR 219

RESULT 8
ID Q84156 PRELIMINARY: PRT: 383 AA.
AC Q84156:
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PHOSPHOLIPASE D ENDONUCLEASE SUPERFAMILY.
GN CR154.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
ON NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/DW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusev R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
DR EMBL: AE001289; AAC67745.1;
DR InterPro: IPR001736; PUD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR Endonuclease; Complete proteome.
SQ
SEQUENCE 383 AA; 43383 MW; 48F73A163553D2FA CRC64;

Query Match 68.1%; Score 32; DB 16; Length 383;
Best Local Similarity 54.5%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 18 SRSRHSPEKR 28

RESULT 9
ID Q9DA19 PRELIMINARY: PRT: 450 AA.
AC Q9DA19:

```

```

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schirml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima Y., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ringwald M., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Welter C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawaji H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK005900; BAB24307.1;
DR MGD: MGI:1914185; 1700023B02RIK.
SQ
SEQUENCE 286 AA; 32818 MW; 068003C5E894827B CRC64;

Query Match 68.1%; Score 32; DB 11; Length 286;
Best Local Similarity 54.5%; Pred. No. 28;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 209 SRSRHSPEKR 219

RESULT 8
ID Q84156 PRELIMINARY: PRT: 383 AA.
AC Q84156:
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE PHOSPHOLIPASE D ENDONUCLEASE SUPERFAMILY.
GN CR154.
OS Chlamydia trachomatis.
OC Bacteria; Chlamydiales; Chlamydiaceae; Chlamydia.
ON NCBI_TaxID=813;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=D/DW-3/CX;
RX MEDLINE=99000809; PubMed=9784136;
RA Stephens R.S., Kalman S., Lammel C.J., Fan J., Marathe R., Aravind L.,
RA Mitchell W.P., Olinger L., Tatusev R.L., Zhao Q., Koonin E.V.,
RA Davis R.W.;
RT "Genome sequence of an obligate intracellular pathogen of humans:
RT Chlamydia trachomatis.";
RL Science 282:754-759(1998).
DR EMBL: AE001289; AAC67745.1;
DR InterPro: IPR001736; PUD.
DR Pfam: PF00614; PLDC; 2.
DR SMART: SM00155; PLDC; 2.
DR Endonuclease; Complete proteome.
SQ
SEQUENCE 383 AA; 43383 MW; 48F73A163553D2FA CRC64;

Query Match 68.1%; Score 32; DB 16; Length 383;
Best Local Similarity 54.5%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 18 SRSRHSPEKR 28

RESULT 9
ID Q9DA19 PRELIMINARY: PRT: 450 AA.
AC Q9DA19:

```

```

DT 01-JUN-2001 (TREMBlrel. 17, Created)
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE 1700023B02RIK.
GN 1700023B02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamamaki I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleischmann W., Gaasterland T., Glass C., King B., Kochia H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamita M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Kombaris P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
RA Wyszynski B., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL: AK006260; BAB24488.1;
DR MGD: MGI:1914185; 1700023B02RIK.
DR InterPro: IPR000345; Cytochrome_C; UNKNOWN_1.
DR PROSITE: PS00190; CYTOCHROME_C; UNKNOWN_1.
SQ SEQUENCE 450 AA; 51852 MW; F32F11BE6D64A6AC CRC64;

Query Match 68.1%; Score 32; DB 11; Length 450;
Best Local Similarity 54.5%; Pred. No. 45;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 373 SRSHRSPEKK 383

RESULT 10
O9XOW5 PRELIMINARY; PRT; 648 AA.
AC O9XOW5;
DT 01-NOV-1999 (TREMBlrel. 12, Created)
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE ATP-DEPENDENT DNA HELICASE.
GN TM1238.
OS Thermotoga maritima.
OC Bacteria; Thermotogales; Thermotoga.
OX NCBI_TaxID=2336;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSB8 / DSM 3109;
RX MEDLINE=99287316; PubMed=10360571;
RA Nelson K.E., Clayton R.A., Gill S.R., Gwinn M.L., Dodson R.J.,
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,
RA Stewart A.M., Sutton M.D., Pratt M.S., Phillips C.A., Richardson D.,
RA Heibelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;
RT "Evidence for lateral gene transfer between Archaea and Bacteria from
RT genome sequence of Thermotoga maritima."

```

```

RL Nature 399:323-329(1999).
DR EMBL: AE001779; RAD36313.1;
DR HSSP: P56255; 1PJR.
DR TIGR: TM1238;
DR InterPro: IPR000212; UvrD-helicase.
DR Pfam: PF00580; UvrD_helicase; 1.
KW Helicase; Complete proteome.
SQ SEQUENCE 648 AA; 75842 MW; 4274FEE319250280 CRC64;

Query Match 68.1%; Score 32; DB 16; Length 648;
Best Local Similarity 62.5%; Pred. No. 66;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RXHXSMEXR 9
DB 370 RSHSHSLE 377

RESULT 11
O9JN22 PRELIMINARY; PRT; 165 AA.
AC O9JN22;
DT 01-OCT-2000 (TREMBlrel. 15, Created)
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE YCG.
GN YCG.
OS Agrobacterium tumefaciens.
OC Plasmid TI.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
OX NCBI_TaxID=358;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=96312368; PubMed=8763954;
RA Alt-Worde J., Stryker J.L., Fugua C., Li P.L., Farrand S.K.,
RA Winans S.C.;
RT "The conjugal transfer system of Agrobacterium tumefaciens octopline-
RT type T1 plasmids is closely related to the transfer system of an IncP
RT plasmid and distantly related to T1 plasmid vir genes.";
RL T. Bacteriol. 178:4248-4257(1996).
RN [2]
RP SEQUENCE FROM N.A.
RA Winans S.C., Zhu J., Oger P.M., Schrammeyer B., Hooykaas P.J.,
RA Farrand S.K.;
RT "Octopline-type T1 plasmid sequence.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF242881; AAF77134.1;
KW Plasmid.
SQ SEQUENCE 165 AA; 18019 MW; 4A3499B6E7696FFC CRC64;

Query Match 66.0%; Score 31; DB 2; Length 165;
Best Local Similarity 54.5%; Pred. No. 26;
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11
DB 116 SRSHSHMAPR 126

RESULT 12
O9JUL6 PRELIMINARY; PRT; 756 AA.
AC O9JUL6;
DT 01-MAY-2000 (TREMBlrel. 13, Created)
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)
DE EG:BACRA2117.8 PROTEIN (CG14625 PROTEIN).
GN EG:BACRA2117.8 OR CG14625.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

```



OC Pterygota: Neoptera: Endopterygota; Diptera: Brachycera; Muscomorpha;  
OC Ephydroidea: Drosophilidae: Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Palazon D., Nelson K.A., Nixon K., Nusskern D.R., Pacib J.M.,  
RA Palazon M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-OREGON-R;  
RA Papadimitrakaki F., Spanos L., Bolshakov V., Siden-Kiamos I., Louis C.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003420; AAF45592.1; ALT\_INIT.  
DR EMBL; AL121806; CAB65885.1;  
DR FLYBase; FBgn040358; EG:BACR42117.8.  
DR InterPro; IPR000873; AMP-binding.  
DR PROSITE; PS00455; AMP\_BINDING; UNKNOWN\_1.  
KW Hypothetical protein.  
SQ SEQUENCE 756 AA; 84285 MW; 06CC714953D1D4DD CRC64;

Query Match 66.0%; Score 31; DB 5; Length 756;  
Best Local Similarity 54.5%; Pred. No. 1.3e+02;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHSMEKR 11  
DB 373 SHSHSEKR 383

RESULT 13  
O9V731 PRELIMINARY; PRT; 800 AA.  
AC O9V731;  
DT 01-MAY-2000 (TReMBLrel. 13, Created)

DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)  
DT 01-OCT-2001 (TReMBLrel. 18, Last annotation update)  
DE CG12869 PROTEIN.  
GN CG12869.  
OS Drosophila melanogaster (fruit fly).  
OC Eukaryota: Metazoa: Arthropoda: Tracheata: Hexapoda: Insecta;  
OC Pterygota: Neoptera: Endopterygota; Diptera: Brachycera; Muscomorpha;  
OC Ephydroidea: Drosophilidae: Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE-20196006; PubMed-10731132;  
RA Adams M.D., Celnik S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotler P.,  
RA Burris K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodra C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Palazon D., Nelson K.A., Nixon K., Nusskern D.R., Pacib J.M.,  
RA Palazon M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-OREGON-R;  
RA Papadimitrakaki F., Spanos L., Bolshakov V., Siden-Kiamos I., Louis C.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AE003420; AAF45592.1; ALT\_INIT.  
DR EMBL; AL121806; CAB65885.1;  
DR FLYBase; FBgn040358; EG:BACR42117.8.  
DR InterPro; IPR000873; AMP-binding.  
DR PROSITE; PS00455; AMP\_BINDING; UNKNOWN\_1.  
KW Hydroxylase.  
SQ SEQUENCE 800 AA; 88490 MW; AB027AB311B6311D CRC64;

Query Match 66.0%; Score 31; DB 5; Length 800;  
Best Local Similarity 62.5%; Pred. No. 1.3e+02;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 HXHSMEKR 11  
DB 448 HSHSMDMR 455

RESULT 14

09W1T8  
ID 09W1T8 PRELIMINARY; PRT; 905 AA.  
AC 09W1T8;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CG13546 PROTEIN.  
GN CG13546.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Gelinkler S.E., Holt R.A., Evans C.A., Gockyne J.D.,  
RA Amaratides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Abdayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.T., Benos P.V., Bertram B.F., Bhandari D., Bolshakov S.,  
RA Borrova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke S., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durlin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwam C.,  
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kethum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laslo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Matzel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mervulov G., Mlshina N.V., Mobarly C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacelj J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reibert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodruff T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RT Science 287:2185-2195(2000).  
RL EMBL: AE003459; AAF46965.1; -;  
DR FlyBase: FBgn0034830; CG13546.  
DR InterPro: IPR003961; FN\_III.  
DR InterPro: IPR003598; IG\_c2.  
DR InterPro: IPR003600; IG\_c2.  
DR InterPro: IPR003006; IG\_MHC.  
DR Pfam: PF00041; fn3.1.  
DR Pfam: PF00047; Ig\_3.  
DR SMART: SM00060; FN3.1.  
DR SMART: SM00408; IGc2.1.  
DR SMART: SM00410; IG\_1ike.1.  
KW Immunoglobulin domain.  
SQ SEQUENCE 905 AA; 96704 MW; F581EC72A2CC8840 CRC64;

Query Match 66.0%; Score 31; DB 5; Length 905;  
Best Local Similarity 50.0%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 RXHSHMEXR 11  
DB 315 RKSHGMOOR 324  
RESULT 15  
ID 09Bj68 PRELIMINARY; PRT; 147 AA.  
AC 09Bj68;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE WINGLESS (FRAGMENT).  
GN WG.  
OS Blackburnia rupicola.  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Coleoptera; Adephaga; Carabidae;  
OC Carabidae Conjuratae; Blackburnia.  
OX NCBI\_TaxID=155387;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Cryan J.R., Liebert J.K., Fetzner J.W., Jr., Whitting M.F.;  
RT "Evaluation of relationships within the endemic Hawaiian Platynini  
RT (Coleoptera: Carabidae) based on molecular and morphological  
RT evidence.";  
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF337627; AAK28269.1; -;  
DR InterPro: IPR000970; Wnt1.  
DR Pfam: PF00110; wntc.2.  
DR SMART: SM00097; WNT1.1.  
FT NON\_TER 1  
FT NON\_TER 147  
SQ SEQUENCE 147 AA; 16318 MW; 10771C5BE3C8F128 CRC64;

Query Match 63.8%; Score 30; DB 5; Length 147;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
OY 2 RXHSHMEXR 11  
DB 14 RVHLHSIKDR 23

Search completed: September 13, 2002, 09:29:28  
Job time: 1073 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:49 ; Search time 80.21 Seconds  
(without alignments)  
5.310 Million cell updates/sec

Title: US-09-446-543a-74  
Perfect score: 47  
Sequence: 1 SRXHXSMEXR 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	91.5	83	1	PRRP_RAT
2	42	89.4	87	1	PRRP_HUMAN
3	42	89.4	98	1	PRRP_BOVIN
4	34	72.3	672	1	HYFB_ECOLI
5	33	70.2	497	1	YY80_MYCTU
6	32	68.1	452	1	YNH4_YEAST
7	32	68.1	460	1	VP41_BPAPS
8	31	66.0	277	1	MCRA_ECOLI
9	31	66.0	282	1	TR2H_HUMAN
10	31	66.0	495	1	HSE1_ARATH
11	30	63.8	333	1	ILVC_MYCAV
12	30	63.8	333	1	ILVC_MYCTU
13	30	63.8	340	1	ILVC_CAMJE
14	30	63.8	772	1	SMN3_RAT
15	30	63.8	1224	1	MSN5_YEAST
16	29	61.7	145	1	OAZ_ONCVO
17	29	61.7	153	1	IEX1_MOUSE
18	29	61.7	219	1	LOL2_MOUSE
19	29	61.7	405	1	IPSP_MOUSE
20	29	61.7	462	1	US29_HCMVA
21	29	61.7	513	1	RFP_HUMAN
22	29	61.7	522	1	RFP_MOUSE
23	29	61.7	774	1	LOL2_HUMAN
24	29	61.7	815	1	YQ07_CAEEL
25	29	61.7	2869	1	RBPL_PLABV
26	28	59.6	277	1	TRFB_DROME
27	28	59.6	319	1	TRXB_BUCAP
28	28	59.6	331	1	UL38_HCMVA
29	28	59.6	338	1	ILVC_CORGL
30	28	59.6	340	1	ILVC_STRTR
31	28	59.6	363	1	ODPA_SMIMA
32	28	59.6	388	1	ODPT_HUMAN
33	28	59.6	389	1	ODPA_ARATH

34	28	59.6	389	1	ODPA_PIG	P29804 sus scrofa
35	28	59.6	390	1	ODPA_HUMAN	P08559 homo sapien
36	28	59.6	390	1	ODPA_MOUSE	P35468 mus musculus
37	28	59.6	390	1	ODPA_RAT	P26284 rattus norv
38	28	59.6	391	1	ODPA_SOLTU	P52903 solanum tub
39	28	59.6	391	1	ODPT_MOUSE	P35487 mus musculus
40	28	59.6	391	1	ODPT_RAT	P26284 rattus norv
41	28	59.6	397	1	ODPA_CAEEL	P52899 caenorhabd
42	28	59.6	397	1	ODPA_PEA	P52902 pisum sativ
43	28	59.6	507	1	C7C3_ARATH	P64638 arabidopsis
44	28	59.6	521	1	IMDH_CHLVI	O50316 chlorella
45	28	59.6	577	1	STR_ECOLI	P11875 escherichia

## ALIGNMENTS

RESULT 1

ID	PRRP_RAT	STANDARD	PRT	83 AA
AC	P81278:			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].			
DE	PRP:			
GN	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96268781; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kikada C., Masuo Y., Amano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RA	"A prolactin-releasing peptide in the brain.";			
RA	Nature 393:272-276(1998).			
RL	[2]			
RN	TISSUE SPECIFICITY.			
RP	PubMed=10498338;			
RX	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kikada C., Kurokawa T., Nishimura O., Onda H.;			
RA	"Tissue distribution of prolactin-releasing peptide (PRP) and its receptor.";			
RL	Regul. Pept. 83:1-10(1999).			
CC	-1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.			
CC	-1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL: AB015418; BAA29026.1;			
KW	Hormone; Amidation; Signal; Cleavage on pair of basic residues.			
FT	1			
FT	PEPTIDE 1			
FT	PEPTIDE 22			
FT	PEPTIDE 33			
FT	MOD RES 52			
SO	SEQUENCE 83 AA; 9215 MW; DOC75A264EEB4F29 CRC64;			

Query Match	91.5%;	Score 43;	DB 1;	Length 83;
Best Local Similarity	72.7%;	Pred. NO. 0.0082;		
Matches	8;	Conservative	0;	Mismatches 3;
			Indels	0;
			Gaps	0;

```
QY      1 SRXHXHSMEXR 11
          || | ||| |
Db      22 SRAHQHSMETR 32
```

RESULT	2	
PRRP_HUMAN		
ID	PRRP_HUMAN	STANDARD;
AC	R01377.	PRT; 87 AA.

DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, last annotation update)  
DE Prolactin-releasing peptide precursor (PrPP) (Prolactin-releasing hormone) [contains: Prolactin-releasing peptide PrnPl; Prolactin-releasing peptide PrnPl20].  
DE Prn.

0C Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
0C Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
0X NCBI\_TaxID=9606;

RP SEQUENCE FROM N.A.  
RC TISSUE-Brain;  
RX MEDLINE-98268781; PubMed-9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,  
KA Kitadawa T., Masuo Y., Asano T., Matsuno H., Sekiguchi M.,  
RT Nishimura O., Onda H., Fujino M.,  
RT "A prolactin-releasing peptide in the brain.",  
PM Nature 393:272-276(1998).

RP TISSUE SPECIFICITY.  
RX Pubmed-10498338;

RA Sengichiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,  
 RA Sumino Y., Fujino M.:  
 RT "tissue distribution of prolactin-releasing peptide (PRP) and its  
 RT receptor".  
 RL Regul. Pept. 83:1-10(1999).  
 CC -1 FUNCTION: Stimulates prolactin (PRL) release and regulates the  
 CC expression of prolactin through its receptor GPR10. May stimulate  
 CC lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/sib.ch>).  
CC or send an email to [license@sib.ch](mailto:license@sib.ch).

DR	EMBL: AB015419; BAA29027.1; -
DR	HIM: 602663; -
KW	hormone; Amidation; Signal.
FT	PEPTIDE 23 53
FT	PEPTIDE 23 53
FT	PEPTIDE 34 53
FT	MOD_RES 53 53
SO	SEQUENCE 87 AA; 9639 MW; 229642F3F50CF981B CCR64;

```

Query Match
Best Local Similarity
Matches      8; Conserv
QY          1 SRXHXSMEXR 11
            11 1 1111 1
Db          23 SRTHRHSMETR 33

```

RESULT	3	
PRRP_BOVIN		
ID_PRRP_BOVIN	STANDARD;	PRT;
00124		98 AA

DT 30-MAY-2000 (rel. 39, Created)  
DT 30-MAY-2000 (rel. 39, Last sequence update)  
DT 01-MAR-2002 (rel. 41, Last annotation update)  
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [contains: Prolactin-releasing peptide PrRP3; Prolactin-releasing peptide PrRP20].

05 *Bos taurus* (Bovine).  
 0C Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:  
 0C Mammalia: Eutheria: Cetartiodactyla: Ruminantia: Pecora: Bovidae:  
 0C Bovidae: Bovinae: Bos.  
 0X NCBI\_TaxID=9913;

RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
RC TISSUE=Brain;  
RA MEDLINE=98268781; PubMed=9607765;  
RX Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,  
RA Kikada C., Masuo Y., Asano T., Matsunoto H., Sekiguchi M.,  
RA Kurokawa T., Nishimura O., Onda H., Fujino M.,  
RT "A prolactin-releasing peptide in the brain.,"  
PL Nature 393:272-276(1998).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the  
CC expression of prolactin through its receptor GPR10. May stimulate  
CC lactotrophs directly to secrete PRL.  
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on lists  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AB015417; BAA29025.1; -  
KM Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT	PEPTIDE	23	53	PROLACTIN-RELEASING PEPTIDE PRRP21.
FT	PEPTIDE	33	53 <td>PROLACTIN-RELEASING PEPTIDE PRRP20.</td>	PROLACTIN-RELEASING PEPTIDE PRRP20.
FT	MOD. RES.	53	53 <td>AMIDATION (G-54 PROVIDE AMIDE GROUP).</td>	AMIDATION (G-54 PROVIDE AMIDE GROUP).
SO	SEQUENCE	98 AA;	10544 MW;	08AC35A13B0FA908 CRC64;
Query Match	Similarity	89.48;	Score 42;	DB 1; Length 98;
Best Local	Similarity	72.78;	Prod. No. 0.016;	
Matches	8; Conservative	0;	Mismatches	3; Indels 0; Gaps 0;

```

QY      1 SRXHXHSMEXR 11
          || | || | |
Db      23 SRAHQHSMETR 33

```

RESULT	4
HYFB_ECOLI	
ID	HYFB_ECOLI
STANDARD;	
PRT;	672 AA

DT	01-NOV-1991 (Rel. 20, Created)
DT	01-NOV-1997 (Rel. 35, last sequence update)
DT	16-OCT-2001 (Rel. 40, last annotation update)
DE	Hydrogenase-4 component B (EC 1.-.-.-).

OS  
OC  
OC  
OC  
OX  
RN  
RP  
RC

HEB ON D2402.  
*Escherichia coli*,  
Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
*Escherichia*.  
NCBI\_TaxID=562;  
(1)  
SEQUENCE FROM N.A., AND CHARACTERIZATION.  
STRAIN-K12;

RA Andrews S.C., McClay J., Ambler A., Quail M., Berks B.C., Guest J.R.;  
 RN Submitted (Oct-1996) to the EMBL/GenBank/DBJ databases.  
 RP [2]  
 RC SEQUENCE FROM N.A.  
 RA STRAIN-K12 / MG1655;  
 RX MEDLINE-97426617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-97349980; PubMed-9205837;  
 RA Yamamoto Y., Alba H., Baba T., Hayashi K., Inada T., Isono K.,  
 RA Itoh T., Kimura S., Kitagawa M., Makino K., Miki T., Mitsuhashi N.,  
 RA Mizobuchi K., Mori H., Nakade S., Nakamura Y., Nishimoto H.,  
 RA Oshima T., Oyama S., Saito N., Sempel G., Satoh Y., Sivasubram S.,  
 RA Tagami H., Takahashi H., Takeda J., Takemoto K., Uehara K., Wada C.,  
 RA Yamagata S., Horiuchi T.;  
 RT "Construction of a contiguous 874-kb sequence of the *Escherichia coli*  
 RT - K12 genome corresponding to 50.0-68.8 min on the linkage map and  
 RT analysis of its sequence features.";  
 RL DNA Res. 4:91-113(1997).  
 RN [4]  
 RP SEQUENCE FROM 1-34 FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-91202105; PubMed-2016588;  
 RA Andrews S.C., Harrison P.M., Guest J.R.;  
 RT "A molecular analysis of the 53.3 minute region of the *Escherichia*  
 RT coli linkage map.";  
 RL J. Gen. Microbiol. 137:361-367(1991).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (By similarity).  
 CC -1- SIMILARITY: TO POLYPEPTIDE 5 OF THE NADH-UBIQUINOL OXIDOREDUCTASE  
 CC OF CHLOROPLASTS OR MITOCHONDRIA.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M63654; AAB88564.1; -;  
 DR EMBL: AE000335; AAC75535.1; -;  
 DR EMBL: D90876; BAA16360.1; -;  
 DR EMBL: D90877; BAA16370.1; -;  
 DR PIR: D49749; D49749;  
 DR EcoGene: EG11282; hvfb.  
 DR InterPro: IPR001750; Oxidored\_q1.  
 DR Pfam: PF00361; oxidored\_q1; 1.  
 DR OXidoreductase; Transmembrane; Inner membrane; Complete proteome.  
 KW  
 FT TRANSMEM 6 POTENTIAL.  
 FT TRANSMEM 31 POTENTIAL.  
 FT TRANSMEM 80 POTENTIAL.  
 FT TRANSMEM 120 POTENTIAL.  
 FT TRANSMEM 140 POTENTIAL.  
 FT TRANSMEM 165 POTENTIAL.  
 FT TRANSMEM 200 POTENTIAL.  
 FT TRANSMEM 220 POTENTIAL.  
 FT TRANSMEM 243 POTENTIAL.  
 FT TRANSMEM 263 POTENTIAL.  
 FT TRANSMEM 273 POTENTIAL.  
 FT TRANSMEM 293 POTENTIAL.  
 FT TRANSMEM 312 POTENTIAL.  
 FT TRANSMEM 332 POTENTIAL.  
 FT TRANSMEM 343 POTENTIAL.  
 FT TRANSMEM 363 POTENTIAL.  
 FT TRANSMEM 385 POTENTIAL.  
 FT TRANSMEM 405 POTENTIAL.  
 FT TRANSMEM 428 POTENTIAL.  
 FT TRANSMEM 448 POTENTIAL.  
 FT TRANSMEM 475 POTENTIAL.  
 FT TRANSMEM 495 POTENTIAL.  
 FT TRANSMEM 505 POTENTIAL.  
 FT TRANSMEM 525 POTENTIAL.  
 FT TRANSMEM 532 POTENTIAL.  
 FT TRANSMEM 552 POTENTIAL.  
 FT TRANSMEM 652 POTENTIAL.  
 FT TRANSMEM 672 POTENTIAL.  
 SQ SEQUENCE 672 AA; 72582 MW; 3CA8C2BD25E54B1B CRC64;

Query Match 72.3%; Score 34; DB 1; Length 672;  
 Best Local Similarity 66.7%; Pred. No. 6.7;  
 Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
 Oy 1 SRXHXSM 9  
 Db 364 SRLHDM 372  
 RESULT 5  
 ID Y80\_MYCTU STANDARD; PRT; 497 AA.  
 AC 006343;  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 53.3 kDa protein RV3480C.  
 GN RV3480C OR MT3584 OR MYCY13E12.33C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE-98295987; PubMed-9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnter T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock R., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Stunson J.E., Taylor K., Whitehead S., Barrall B.G.;  
 RT "Deciphering the biology of *Mycobacterium tuberculosis* from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 RA Kolony J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 RA Delcher A., Utterback T., Weidman J., Khouli H., Gill J., Mikula A.,  
 RA Bishai W.;  
 RT "Whole genome comparison of *Mycobacterium tuberculosis* clinical and  
 RT laboratory strains.";  
 RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: BELONGS TO THE UPF0089 FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: Z95390; CAB08703.1; -;  
 DR EMBL: AE007161; AAK47943.1; ALT\_INIT.  
 DR TIGR: MT3584; -;  
 DR Tuberculist: RV3480C; -;  
 DR InterPro: IPR004255; UPF0089.  
 DR Pfam: PF03007; UPF0089; 1.  
 KW Hypothetical protein; Complete proteome.  
 FT CONFLICT 427 427 L -> V (IN REF. 2).  
 SQ SEQUENCE 497 AA; 53314 MW; 90D661E32B5658C2 CRC64;

Query Match 70.2%; Score 33; DB 1; Length 497;  
 Best Local Similarity 45.3%; Pred. No. 7.8;

Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRHXHSMEXR 11  
 DB 108 SRHSHALDLR 118

RESULT 6  
 ID YNH4\_YEAST STANDARD; PRT; 452 AA.  
 AC P32047;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DE 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Hypothetical 49.2 kDa protein IMP4-MSK1 intergenic region.  
 GN YNL074C OR N2359 OR YMK1.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C / FY1679;  
 RX MEDLINE=96267764; PubMed=8701611;  
 RA Pohlmann R., Phillipsen P.;  
 RT "Sequencing a cosmid clone of Saccharomyces cerevisiae XIV  
 RT reveals 12 new open reading frames (ORFs) and an ancient duplication  
 RT of six ORFs."  
 RL Yeast 12:391-402(1996).  
 RN [2]  
 RP SEQUENCE OF 1-217 FROM N.A.  
 RX MEDLINE=91202535; PubMed=2016746;  
 RA Gatti D., Tzagoloff A.;  
 RT "Structure and evolution of a group of related aminoacyl-tRNA  
 RT synthetases."  
 RL J. Mol. Biol. 218:557-568(1991).  
 CC -1 SIMILARITY: NO YEAST YH135C.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X86470; CAA60186.1; -  
 DR EMBL: 271350; CAA95948.1; -  
 DR EMBL: X57360; -; NOT\_ANNOTATED\_CDS.  
 DR PIR: S27290; S27290.  
 DR SGD: S0003018; YNL074C.  
 KW Hypothetical protein.  
 FT DOMAIN 211 214 POLY-TIR.  
 FT DOMAIN 271 274 POLY-SER.  
 SO SEQUENCE 452 AA; 49210 MW; 9C4FB0D38D6963E CRC64;

Query Match 68.1%; Score 32; DB 1; Length 452;  
 Best Local Similarity 54.5%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRHXHSMEXR 11  
 DB 189 SRSHSHSAATR 199

RESULT 7  
 ID VP41\_BPAPS STANDARD; PRT; 460 AA.  
 AC Q9T107;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Putative protein P41.  
 GN 41.  
 OS Bacteriophage APSE-1.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae.  
 OX NCBI\_TaxID=106199;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99420383; PubMed=10489345;  
 RA van der Wilk F., Dullemeers A.M., Verbeek M., van den Heuvel J.F.J.M.;  
 RT "Isolation and characterization of APSE-1, a bacteriophage infecting  
 RT the secondary endosymbiont of acyrthosiphon pisum."  
 RL Virology 262:104-113(1999).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF157835; AAF03984.1; -  
 KW Hypothetical protein  
 SO SEQUENCE 460 AA; 51920 MW; 9572842E1FFDDE31 CRC64;

Query Match 68.1%; Score 32; DB 1; Length 460;  
 Best Local Similarity 55.6%; Pred. No. 12;  
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 SRHXHSMEXR 9  
 DB 259 ARAHYHAME 267

RESULT 8  
 ID MCRA\_ECOLI STANDARD; PRT; 277 AA.  
 AC P24200;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE 5-methylcytosine-specific restriction enzyme A (EC 3.1.21.-)  
 DE (EcoKMcrA).  
 GN MCRA OR RGLA OR B1159.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=92041637; PubMed=1938927;  
 RA Hlom K.J., Sedgwick S.G.;  
 RT "Cloning and structural characterization of the mcrA locus of  
 RT Escherichia coli."  
 RL J. Bacteriol. 173:7368-7373(1991).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ramalingam R., Prasad R., Shivapriya R., Dharmalingam K.;  
 RT "Molecular cloning and sequencing of mcrA locus and identification of  
 RT MCR protein in Escherichia coli."  
 RL J. Biosci. 17:217-232(1992).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12."  
 RL Science 277:1453-1474(1997).  
 RN [4]

RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-97061202; PubMed-8905232;  
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A.,  
 RA Iemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K.,  
 RA Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K.,  
 RA Mori H., Motomura K., Nakamura Y., Nishimoto H., Nishio Y., Saito N.,  
 RA Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y.,  
 RA Yano M., Horinouchi T.;  
 RT "A 718-kb DNA sequence of the *Escherichia coli* K-12 genome  
 corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 CC -1- FUNCTION: RESTRICTION OF 5-METHYL AND 5-HYDROXYMETHYLCYTOSINES AT  
 CC THE SPECIFIC DNA SEQUENCE C(M)CGG.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M76667; AAA68481.1; -;  
 DR EMBL: Z19104; CA979520.1; -;  
 DR EMBL: AE000215; AAC74243.1; -;  
 DR EMBL: D90750; BAA35995.1; -;  
 DR PIR: A41424; A41424.  
 DR REBASE: 2832; ECKMGFA.  
 DR Ecogene: EG10573; mcrA.  
 DR InterPro: IPR002711; HNH.  
 DR InterPro: IPR003615; HNH\_Luc.  
 DR Pfam: PF01844; HNH\_1.  
 DR SMART: SM00507; HNHc\_1.  
 KW Hydrolyase: Endonuclease; Nuclease; Restriction system;  
 KM Complete proteome.  
 FT CONFLICT 214 A -> R (IN REF. 1).  
 FT SEQUENCE 277 AA; 31390 MW; E5F2627DFEDEC402 CRC64;  
 SQ

Query Match 66.0%; Score 31; DB 1; Length 277;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 RXHXHSMEXR 11  
 I I I I I I  
 DB 69 RKHMHSIDER 78

RESULT 9  
 TR2H\_HUMAN STANDARD; PRT; 282 AA.  
 AC Q13595;  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Transformer-2 protein homolog (TFA-2 alpha).  
 GN TRA2A.  
 OS Homo sapiens (Human).  
 OC Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;  
 OC Mammalia: Eutheria: Primates: Catarrhini: Hominoidea: Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96392356; PubMed-8799144;  
 RA Dauwalder B., Amaya-Manzanares F., Mattox M.;  
 RT "A human homologue of the *Drosophila* sex determination factor  
 RT transformer-2 has conserved splicing regulatory functions.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 93:9004-9009(1996).  
 CC -1- FUNCTION: MAY INTERACT WITH SPECIFIC PRE-MRNAS AND AFFECT THEIR  
 CC SPLICING PATTERNS.  
 CC -1- ALTERNATIVE PRODUCTS: 2 ISOFORMS: A LONG FORM (SHOWN HERE) AND A  
 CC SHORT FORM; ARE PRODUCED BY ALTERNATIVE SPLICING.

CC -1- PTM: EXTENSIVELY PHOSPHORYLATED ON SERINE RESIDUES IN THE RS  
 CC DOMAIN (BY SIMILARITY).  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -1- SIMILARITY: BELONGS TO THE SR FAMILY OF SPLICING FACTORS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U53209; AAC50658.1; -;  
 DR MIM: 602718; -;  
 DR InterPro: IPR000504; RRM.  
 DR Pfam: PF00076; rrm; 1.  
 DR SMART: SM00360; RRM; 1.  
 DR PROSITE: PS50102; RRM; 1.  
 DR PROSITE: PS00030; RRM\_RNP\_1; 1.  
 KW RNA-binding; mRNA splicing; Alternative splicing; Phosphorylation.  
 FT DOMAIN 30 112 ARG/SER-RICH (RS1 DOMAIN).  
 FT DOMAIN 119 197 RNA-BINDING (RRM).  
 FT DOMAIN 198 225 LINKER.  
 FT VARSPLIC 226 282 ARG/SER-RICH (RS2 DOMAIN).  
 FT VARSPLIC 1 101 MISSING (IN SHORT ISOFORM).  
 FT VARSPLIC 214 214 H -> Q (IN SHORT ISOFORM).  
 FT VARSPLIC 215 282 MISSING (IN SHORT ISOFORM).  
 SQ SEQUENCE 282 AA; 32688 MW; EDB5ABF7BEA023FD CRC64;

Query Match 66.0%; Score 31; DB 1; Length 282;  
 Best Local Similarity 54.5%; Pred. No. 11;  
 Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 SRHXHSMEXR 11  
 I I I I I I  
 DB 71 SRSHSHRR 81

RESULT 10  
 HSFL\_ARATH STANDARD; PRT; 495 AA.  
 ID HSFL\_ARATH  
 AC P41151; O23615;  
 DT 01-FEB-1995 (Rel. 31, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Heat shock factor protein 1 (HSF 1) (Heat shock transcription factor  
 DE 1) (HSTF 1).  
 GN HSF1 OR AT4G1750 OR DL4910C.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta;  
 OC Spermatophyta: Magnoliopsida: eudicotyledons: core eudicots: Rosidae;  
 OC eustosids II: Brassicales: Brassicaceae: Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-95036006; PubMed-7948881;  
 RA Huebel A., Schoeffl F.;  
 RT "Arabidopsis heat shock factor: Isolation and characterization of the  
 RT gene and the recombinant protein.";  
 RL Plant Mol. Biol. 26:353-362(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX STRAIN-CV, COLUMBIA;  
 RX MEDLINE-96121113; PubMed-9461215;  
 RA Bevan M., Bancroft I., Bent E., Love K., Goodman H., Dean C.,  
 RA Beggs R., Dirkse W., van Slaveren M., Stiekema W., Drost L.,  
 RA Ridley P., Hudson S.A., Patel K., Murphy G., Piffanelli P., Wedler H.,  
 RA Wedler E., Wambutt R., Weitzenecker T., Pohl T.M., Terry N.,  
 RA Gleen J., Villarroel R., De Clerck R., van Montagu M., Lechany A.,  
 RA Auborg S., Gy I., Kreis M., Lao N., Kavanagh T., Hempel S., Kotler P.,  
 RA Enlita K.D., Rieger W., Schaefer M., Funk B., Mueller-Auer S.,

RA Silvey M., James R., Montfort A., Pons A., Pulgomenech P., Douka A.,  
RA Vouletou E., Milioni D., Hatzopoulos P., Piravandi E., Obermayer B.,  
RA Hilbert H., Duesterhoft A., Moores T., Jones J.D.G., Eneva T.,  
RA Palmer K., Benes V., Reckman S., Ansoerge W., Cooke R., Berger C.,  
RA Delany M., Voet M., Volckaert G., Mewes H.-W., Klosterman S.,  
RA Scheller C., Chalwatzis N.,  
RT Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of  
RT Arabidopsis thaliana." ;  
RL Nature 391:485-488(1998).  
[3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CV COLUMBIA:  
RX MEDLINE-20083488; PubMed-10617198;  
RA Mayer K.F.X., Scheller C., Wandt R., Murphy G., Volckaert G.,  
RA Pohl T., Duesterhoft A., Stiekema W., Entian K.-D., Terry N.,  
RA Harris B., Ansoerge W., Brandt P., Griwall L.A., Rieger M.,  
RA Wehlselgatter M., de Simone V., Obermayer B., Macho R., Mueller M.,  
RA Kreis M., Delany M., Pulgomenech P., Watson M., Schmidtnehl T.,  
RA Reichert B., Portetelle D., Perez-Alonso M., Boutry M., Bancroft I.,  
RA Vos P., Hohnel J., Zimmermann W., Wedler H., Ridley P.,  
RA Langham S.-A., McCullagh B., Bilham L., Roben J.,  
RA Van der Schueren J., Grymoprez B., Chuang Y.-J., Vandenbussche F.,  
RA Braeken M., Weltjens I., Voet M., Bastiaens I., Aert R., Defoor E.,  
RA Weizengger T., Bothe G., Ramsperger U., Hilbert H., Braun M.,  
RA Holzer E., Brandt A., Peters S., Van Staveren M., Dirse W.,  
RA Mooijman P., Klein Lankhorst R., Rose M., Hauf J., Koeltter P.,  
RA Bernier S., Hempel S., Feldpausch M., Lamberth S., Van den Daele H.,  
RA De Keyser A., Buysaert C., Gielen J., Villarroel R., De Clercq R.,  
RA Van Montagu M., Rogers J., Cronin A., Quail M., Bray-Allen S.,  
RA Clark L., Doggett J., Hall S., Kay M., Lennard N., McIay K., Mayes R.,  
RA Petre A., Rajanoram M.A., Lyne M., Benes V., Reckmann S.,  
RA Borova D., Bloecker H., Scharte M., Grimm M., Loehner T.-H.,  
RA Dose S., de Haan M., Maarse A.C., Schaefer M., Mueller-Auer S.,  
RA Gabel C., Fuchs M., Fartmann B., Granderath K., Dauner D., Herzl A.,  
RA Neumann S., Argitlou A., Vitale D., Liguori R., Piravandi E.,  
RA Masencot O., Outley F., Clabaud G., Muendlein A., Felber R.,  
RA Schnabl S., Hiller R., Schmidt W., Lecharny A., Aubourg S.,  
RA Cheidori F., Cooke R., Berger C., Montfort A., Casacuberta E.,  
RA Gibbons T., Weber N., Vandenbol M., Bagues M., Terol J., Torres A.,  
RA Perez-Perez A., Purnelle B., Bent E., Johnson S., Tacon D., Jesse T.,  
RA Heijnen L., Schwarz S., Scholler P., Heber S., Francis P., Blake C.,  
RA Frishman D., Haase D., Lemcke K., Mewes H.-W., Stocker S.,  
RA Zaccaria P., Bevan M., Wilson R.K., de la Bastide M., Habermann K.,  
RA Parnell L., Dedha N., Gnoj L., Schuit K., Huang E., Spiegel L.,  
RA Sekhon M., Murray J., Sheet P., Cordes M., Abu-Threidat J.,  
RA Stoneking T., Kallik J., Graves T., Harmon G., Edwards J.,  
RA Latreille P., Courtney L., Cloud J., Abbott A., Scott K., Johnson D.,  
RA Mink P., Bentley D., Fulton B., Miller N., Greco T., Kemp K.,  
RA Kramer J., Fulton L., Maridis E., Dante M., Pepin K., Hillier L.,  
RA Nelson J., Spieth J., Ryan E., Andrews S., Gelsel C., Layman D.,  
RA Du H., Ali J., Berghoff A., Jones K., Drone K., Colton M., Joshi C.,  
RA Antonou B., Zidanic M., Strong C., Sun H., Lamar B., Jordan C.,  
RA Ma P., Zhong J., Preston R., Vill D., Shekhar M., Matero A., Shah R.,  
RA Swaby I.K., O'Shaughnessy A., Rodriguez M., Hoffman J., Tili S.,  
RA Gravat S., Shohdy N., Hasegawa A., Hamed A., Lohd M., Johnson A.,  
RA Chen E., Maria M., Martensen R., McCombie W.R.;  
RT "Sequence and analysis of chromosome 4 of the plant Arabidopsis  
RT thaliana." ;  
RL Nature 402:769-777(1999).  
CC -1- FUNCTION: DNA-BINDING PROTEIN THAT SPECIFICALLY BINDS HEAT SHOCK  
CC PROMOTER ELEMENTS (HSE) AND ACTIVATES TRANSCRIPTION (BY  
CC SIMILARITY).  
CC -1- SUBUNIT: HOMOTRIMER.  
CC -1- SUBCELLULAR LOCATION: Nuclear.  
CC -1- PTM: EXHIBITS TEMPERATURE-DEPENDENT PHOSPHORYLATION  
CC (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE HSF FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X76167; CAA53761.1; -  
DR EMBL: 297344; CAB10555.1; -  
DR EMBL: AL161547; CAB78778.1; -  
DR PIR: S38873; S38873.  
DR HSP: P22121; 3HSF.  
DR InterPro: IPR000232; HSF\_DNA\_bind.  
DR InterPro: IPR002341; HSF\_ETS.  
DR Pfam: PF00447; HSF\_DNA-Bind; 1.  
DR PRINTS: PR00056; HSFDOMAIN.  
DR ProDom: PD00178; HSF\_DNA\_bind; 1.  
DR SMART: SM00415; HSF; 1.  
DR PROSITE: PS00434; HSF\_DOMAIN; 1.  
KW Transcription regulation; Nuclear protein; DNA-binding; Activator;  
KW Phosphorylation; Heat shock; Multigene family.  
FT DNA\_BIND 50 144  
FT DOMAIN 150 153  
FT DOMAIN 204 207  
FT DOMAIN 454 475  
FT DOMAIN 216 216  
FT CONFLICT 285 285  
FT CONFLICT 336 336  
FT CONFLICT 390 390  
FT CONFLICT 462 473  
SQ SEQUENCE 495 AA; 55744 MW; 2F0DE5525B5682A CRC64;  
  
OY 1 SRXHXSMSE 9  
DB 277 SNSSHSLP 285  
  
RESULT 11  
ITVC MYCAV STANDARD; PRT; 333 AA.  
AC 059500;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ketol-acid reductoisomerase (EC 1.1.1.86) (Acetohydroxy-acid  
DE isomerase) (Alpha-keto-beta-hydroxyacyl reductoisomerase).  
GN ITVC.  
OS Mycobacterium avium.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
OX NCBI\_TaxID-1764;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97080504; PubMed-8921849;  
RA Gusberti L., Cantoni R., de Rossi E., Branzoni M., Riccardi G.;  
RT "Cloning and sequencing of the ilvNC gene cluster from Mycobacterium  
RT avium." ;  
RL Gene 177:83-85(1996).  
CC -1- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)  
CC = (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.  
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS; SECOND STEP.  
CC -1- SIMILARITY: BELONGS TO THE KETOL-ACID REDUCTOISOMERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: L49392; AAB38428.1; -



DR InterPro: IPR000506; Acetylhydrazac\_1somrdctase.  
DR Pfam: PF01450; ILVC: 1.  
KW Oxidoreductase; Branched-chain amino acid biosynthesis; NADP.  
FT ACT\_SITE 105 105 POTENTIAL.  
SQ SEQUENCE 333 AA; 36058 MW; 336AC0BD51563DE2 CRC64;

Query Match 63.88; Score 30; DB 1; Length 333;  
Best Local Similarity 45.58; Pred. No. 22;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
|:|:|:|  
Db 24 SQGHAHSLSLR 34

RESULT 12  
ILVC\_MYCTU STANDARD: PRT: 333 AA.  
AC 053246;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ketol-acid reductoisomerase (EC 1.1.1.86) (Acetylhydroxy-acid  
isomeroeductase) (Alpha-keto-beta-hydroxyacyl reductoisomerase).  
GN ILVC OR RV3001C OR MT3081 OR MTV012.15C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.  
OX NCBI\_TaxID-1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RX MEDLINE-98295987; PubMed-9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eiglmeller K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Baeham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Kirogh A., Mclean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen Mycobacter tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CDC 1551 / Oshkosh;  
RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
RA Peterson J., Deboy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
RA Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
RA Delcher A., Utterback T., Weidman J., Khouri H., Gill J., Mikula A.,  
RA Bishai W.;  
RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
RT laboratory strains."  
RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
CC -1- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)  
CC = (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.  
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS: SECOND STEP.  
CC -1- SIMILARITY: BELONGS TO THE KETO-ACID REDUCTOISOMERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AL021287; CAA16086.1; -  
DR EMBL: AE007128; AAK47410.1; ALU\_INTL.  
DR TIGR: MT3081; -  
DR TubercuList; RV3001C; -

DR InterPro: IPR000506; Acetylhydrazac\_1somrdctase.  
DR Pfam: PF01450; ILVC: 1.  
KW Oxidoreductase; Branched-chain amino acid biosynthesis; NADP;  
FT Complete proteome.  
FT ACT\_SITE 105 105 POTENTIAL.  
SQ SEQUENCE 333 AA; 36090 MW; 49669E3EDACE25EC CRC64;

Query Match 63.88; Score 30; DB 1; Length 333;  
Best Local Similarity 45.58; Pred. No. 22;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
|:|:|:|  
Db 24 SQGHAHSLSLR 34

RESULT 13  
ILVC\_CAMJE STANDARD: PRT: 340 AA.  
AC 09PHN5;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Ketol-acid reductoisomerase (EC 1.1.1.86) (Acetylhydroxy-acid  
isomeroeductase) (Alpha-keto-beta-hydroxyacyl reductoisomerase).  
GN ILVC OR CJ0632.  
OS Campylobacter jejuni.  
OC Bacteria; Proteobacteria; epsilon subdivision; Campylobacter group;  
OC Campylobacter.  
OX NCBI\_TaxID-197;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-NCCTC 11168;  
RX MEDLINE-20150912; PubMed-10688204;  
RA Parkhill J., Wren B.W., Mungall K., Kelley J.M., Churcher C.,  
RA Baeham D., Chillingworth T., Davies R.M., Felwell T., Holroyd S.,  
RA Jagels K., Karlyshay A.V., Moule S., Pallen M.J., Penn C.W.,  
RA Quail M.A., Rajandream M.A., Rutherford K.M., van Vliet A.H.M.,  
RA Whitehead S., Barrell B.G.;  
RT "The genome sequence of the food-borne pathogen Campylobacter jejuni  
RT reveals hypervariable sequences."  
RL Nature 403:665-668(2000).  
CC -1- CATALYTIC ACTIVITY: (R)-2,3-dihydroxy-3-methylbutanoate + NADP(+)  
CC = (S)-2-hydroxy-2-methyl-3-oxobutanoate + NADPH.  
CC -1- PATHWAY: VALINE AND ISOLEUCINE BIOSYNTHESIS: SECOND STEP.  
CC -1- SIMILARITY: BELONGS TO THE KETO-ACID REDUCTOISOMERASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AL139075; CAB75268.1; -  
DR InterPro: IPR000506; Acetylhydrazac\_1somrdctase.  
DR Pfam: PF01450; ILVC: 1.  
KW Oxidoreductase; Branched-chain amino acid biosynthesis; NADP;  
FT Complete proteome.  
FT ACT\_SITE 109 109 POTENTIAL.  
SQ SEQUENCE 340 AA; 36934 MW; DA556ED90C86C780 CRC64;

Query Match 63.88; Score 30; DB 1; Length 340;  
Best Local Similarity 45.58; Pred. No. 22;  
Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
|:|:|:|  
Db 28 SQGHAHSLSLR 38

```

RESULT 14
SM3A_RAT 14
ID SM3A_RAT STANDARD: PRT: 772 AA.
AC 063548:
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Semaphorin 3a precursor (Semaphorin III) (Sema III).
GN SEMA3A.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MSTAR; TISSUE=Brain;
RX MEDLINE=97073089; PubMed=8915837;
RA Giger R.J., Wolfer D.P., De Wit G.M.J., Verhaagen J.;
RT "Anatomy of rat semaphorin III/collapsin-1 mRNA expression and
RT relationship to developing nerve tracts during neuroembryogenesis.";
RL J. Comp. Neurol. 375:378-392(1996).
CC -1 FUNCTION: MAY BE INVOLVED IN GUIDING GROWING AXONS TOWARDS THEIR
CC TARGETS BY FORMING A MOLECULAR BOUNDARY THAT INSTRUCTS AXONS TO
CC ENGAGE IN THE FORMATION OF SPECIFIC NERVE TRACTS. BINDS TO
CC NEUROFILIN.
CC -1 SUBCELLULAR LOCATION: Secreted (by similarity).
CC -1 DEVELOPMENTAL STAGE: AT E-11, EXPRESSION WAS RESTRICTED TO THE
CC OLFACATORY PIT, THE BASAL AND ROSTRAL SURFACE OF THE TELENCEPHALIC
CC VESICLE, THE EYE ANLAGE, THE EPITHELIUM OF RATHKE'S POUCH, AND
CC SOMITES. AT LATER DEVELOPMENTAL STAGES, IT WAS WIDELY DISTRIBUTED
CC IN NEURONAL AS WELL AS IN MESENCHYMAL AND EPITHELIAL STRUCTURES
CC OUTSIDE THE NERVOUS SYSTEM. AFTER BIRTH, MESENCHYMAL LEVELS
CC DECREASED RAPIDLY AND EXPRESSION BECAME RESTRICTED TO SPECIFIC
CC SETS OF NEURONS IN THE CNS. IN THE MATURE CNS, IT IS DETECTABLE IN
CC MITRAL CELLS, NEURONS OF THE ACCESSORY BULB AND CEREBRAL CORTEX,
CC CEREBELLAR PURKINJE CELLS, AS WELL AS A SUBSET OF CRANIAL AND
CC SPINAL MOTORNEURONS.
CC -1 DOMAIN: STRONG BINDING TO NEUROFILIN IS MEDIATED BY THE CARBOXY
CC THIRD OF THE PROTEIN.
CC -1 SIMILARITY: BELONGS TO THE SEMAPHORIN FAMILY.
CC -1 SIMILARITY: CONTAINS 1 SEMA DOMAIN.
CC -1 SIMILARITY: CONTAINS 1 IMMUNOGLOBULIN-LIKE C2-TYPE DOMAIN.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X93286; CAA64607.1; -.
DR InterPro: IPR003599; IG.
DR InterPro: IPR003006; IG_MHC.
DR InterPro: IPR003659; PSI.
DR InterPro: IPR001627; Sema.
DR Pfam: PF00047; Ig_1.
DR Pfam: PF01403; Sema; 1.
DR SMART: SM00409; IG; 1.
DR SMART: SM00423; PSI; 1.
DR Signal: Immunoglobulin domain; Multigene family; Neurogenesis;
KM Developmental protein; Glycoprotein.
FT SIGNAL 1 20
FT CHAIN 1 772
FT DOMAIN 21 772 SEMAPHORIN 3A.
FT DOMAIN 240 538 SEMA.
FT DOMAIN 643 730 IG-LIKE C2-TYPE DOMAIN.
FT DISULFID 728 770 ARG/LYS-RICH (BASIC).
FT CARBOHYD 53 53 BY SIMILARITY.
FT CARBOHYD 125 125 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 591 591 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 772 AA; 88808 MW; 240907812FF9F2D2 CRC64;

```

```

Query Match 63.8%; Score 30; DB 1; Length 772;
Best Local Similarity 62.5%; Pred. No. 55;
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
OY 4 HXHSMEXR 11
DB 577 HGHSLER 584

```

```

RESULT 15
MSNS_YEAST
ID MSNS_YEAST STANDARD: PRT: 1224 AA.
AC P52918:
DT 01-OCT-1996 (Rel. 34, Created)
DT 01-OCT-1996 (Rel. 34, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE MSNS protein.
GN MSNS OR YDR335W OR D9651.5.
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.
OX NCBI_TaxID=4932;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C;
RX Alepuz P.M., Estruch F.;
RT Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=S288C / AB972;
RA Favello A., Andrews S., Brinkman R., Cooper J., Ding H., Du Z.,
RA Johnson D., Johnston L., Langston Y., Latreille P., Le T.,
RA Mardis E., Menezes S., Miller N., Nhan M., Pauley A., Peluso D.,
RA Rifkin L., Riles L., Tatch A., Trevisan E., Vignati D.,
RA Wilcox L., Wohlman P., Vaudin M., Wilson R., Waterston R.;
RL Submitted (MAR-1996) to the EMBL/GenBank/DBJ databases.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X93302; CAA63705.1; -.
DR EMBL: U51032; AAB64771.1; -.
DR SGD: S0002743; MSNS.
DR CONFLICT 886 886 D -> G (IN REF. 1).
FT CONFLICT 1011 1011 S -> R (IN REF. 1).
SQ SEQUENCE 1224 AA; 142117 MW; C47A6767CBA44486 CRC64;

```

```

Query Match 63.8%; Score 30; DB 1; Length 1224;
Best Local Similarity 54.5%; Pred. No. 92;
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
OY 1 SRXHSMEXR 11
DB 1014 SRLNHMEXR 1024

```

Search completed: September 13, 2002, 09:30:50  
Job time: 1140 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:05 ; Search time 172.41 seconds  
(without alignments)  
6.131 Million cell updates/sec

Title: US-09-446-543a-74  
Perfect score: 47  
Sequence: 1 SRXHXSMEXR 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	91.5	83	2 JC7607	prolactin-releasing peptide - human
2	35	74.5	451	2 G01227	hydrogenase-4 comp
3	34	72.3	672	2 A65024	hydrogenase 4 memb
4	34	72.3	672	2 H91046	hydrogenase 4 memb
5	34	72.3	672	2 D85891	hypothetical prote
6	33	70.2	497	2 E70568	hypothetical prote
7	32	68.1	196	2 D84351	probable endonucle
8	32	68.1	452	2 F71550	YMK1 protein - yea
9	32	68.1	452	2 S53906	ATP-dependent DNA
10	32	68.1	648	2 G72279	5-methylcytosine-s
11	31	66.0	277	2 A41424	heat shock transcr
12	31	66.0	495	1 S52641	thioredoxin reduct
13	30	63.8	330	2 H84075	thioredoxin reduct
14	30	63.8	332	1 B70015	ketol-acid reducto
15	30	63.8	333	1 JC5166	probable 11vc prot
16	30	63.8	333	2 D70855	ketol-acid reducto
17	30	63.8	340	2 G81411	hypothetical prote
18	30	63.8	393	2 T28994	hypothetical prote
19	30	63.8	509	2 T67910	hypothetical prote
20	30	63.8	599	2 F82291	conserved hypotet
21	30	63.8	1224	2 S70100	MEN5 protein - yea
22	29	61.7	122	2 G81198	hypothetical prote
23	29	61.7	122	2 G81775	hypothetical prote
24	29	61.7	153	2 S33363	gamma-secalin - mo
25	29	61.7	194	2 A23277	hypothetical prote
26	29	61.7	292	2 T05317	transactin initia
27	29	61.7	307	2 T17104	HHRF4 protein - hu
28	29	61.7	462	1 QOBE04	protein F57B9.10 l
29	29	61.7	467	2 H88493	

30	29	61.7	506	2 S37583	RING finger protei
31	29	61.7	513	1 TVH0RF	ret finger protein
32	29	61.7	569	2 F75381	probable two-compo
33	29	61.7	815	2 T15402	hypothetical prote
34	29	61.7	2829	2 A42771	reticulocyte-bind
35	28	59.6	223	2 A60226	pyruvate dehydroge
36	28	59.6	106	2 G87103	conserved hypotet
37	28	59.6	178	2 T49712	hypothetical prote
38	28	59.6	233	2 F87216	probable TetR-fam1
39	28	59.6	234	2 D70829	probable regulator
40	28	59.6	277	2 S57336	transcription init
41	28	59.6	295	2 T48869	serotonln receptor
42	28	59.6	302	4 J02275	hypothetical prote
43	28	59.6	303	2 T47356	glycosyltransferas
44	28	59.6	303	2 A84114	probable N-acetyl
45	28	59.6	312	2 C71262	

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, A.;  
Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior  
release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PrRP  
A:Introns: 33/1

Query Match 91.5%; Score 43; DB 2; Length 83;  
Best Local Similarity 72.7%; Pred. No. 0.07;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy 1 SRXHXSMEXR 11  
|||  
Db 22 SRAHQSMETR 32

RESULT 2  
G01227  
receptin - human  
C:Species: Homo sapiens (man)  
C:Date: 21-Dec-1996 #sequence\_revision 06-Jun-1997 #text\_change 05-Nov-1999  
C:Accession: G01227  
R:Chao, L.  
submitted to the EMBL Data Library, November 1993  
A:Reference number: G06292  
A:Accession: G01227  
A:Status: preliminary; translated from GR/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-451 <CHA>  
A:Cross-references: EMBL:U03644; NID:9476104; PIDN:AAA17853.1; PID:9476105  
C:Genetics:  
A:Gene: receptin

Query Match 74.5%; Score 35; DB 2; Length 451;  
Best Local Similarity 63.6%; Pred. No. 9.6;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXHSMEXR 11  
|||  
Db 360 SRTHKHSPEKR 370

RESULT 3  
A:Accession: A65024  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 31-Dec-2001  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
. ; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: A65024  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-672 <BLAT>  
A:Cross-references: GB:AE000335; GB:U00096; NID:g1788821; PIDN:AAC75535.1; PID:g1788827;  
A:Experimental source: strain K-12, substrain MG1655  
R:Andrews, S.C.; Harrison, P.M.; Guest, J.R.  
J. Gen. Microbiol. 137, 361-367, 1991  
A:Title: A molecular analysis of the 53.3 minute region of the Escherichia coli linkage  
A:Reference number: A49749; MUID:91202105  
A:Accession: D49749  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-34 <AND>  
A:Cross-references: GB:M37689  
C:Superfamily: formate hydrogenlyase chain 3  
C:Keywords: oxidoreductase

Query Match 72.3%; Score 34; DB 2; Length 672;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 SRXHXHSMEXR 9  
|||  
Db 364 SRTHHDM 372

RESULT 4  
H91046  
hydrogenase 4 membrane subunit [imported] - Escherichia coli (strain O157:H7, substrain  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: H91046  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gesawara, N.; Yasunaga, T.; Kuhara, S.; Shibata, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno  
A:Reference number: A99639; MUID:21156231; PMID:11258796  
A:Accession: H91046  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-672 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA36767.1; PID:g13362814; GSPDB:GN00154  
C:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECs3344  
C:Superfamily: formate hydrogenlyase chain 3

Query Match 72.3%; Score 34; DB 2; Length 672;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
OY 1 SRXHXHSMEXR 9  
|||

Db 364 SRTHHDM 372

RESULT 5  
D85891  
hydrogenase 4 membrane subunit [imported] - Escherichia coli (strain O157:H7, substra  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: D85891  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
11ler, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Dialianta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A65480; MUID:21074935; PMID:11206551  
A:Accession: D85891  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-672 <STO>  
A:Cross-references: GB:AE005174; NID:g12516867; PIDN:AMG57592.1; GSPDB:GN00145; UWGP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: hlyE  
C:Superfamily: formate hydrogenlyase chain 3

Query Match 72.3%; Score 34; DB 2; Length 672;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXHSMEXR 9  
|||  
Db 364 SRTHHDM 372

RESULT 6  
E70568  
hypothetical protein RV3480c - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: E70568  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:98295987  
A:Accession: E70568  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-497 <COB>  
A:Cross-references: GB:Z95390; GB:AL123456; NID:g3261766; PIDN:CAB08703.1; PID:e31607  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV3480c

Query Match 70.2%; Score 33; DB 2; Length 497;  
Best Local Similarity 45.5%; Pred. No. 28;  
Matches 5; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
OY 1 SRXHXHSMEXR 11  
|||  
Db 108 SRHSHADLR 118

RESULT 7  
D84351  
hypothetical protein Vng2006c [imported] - Halobacterium sp. NRC-1  
C:Species: Halobacterium sp. NRC-1  
C:Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001  
C:Accession: D84351  
R:Ng, W.V.; Kennedy, S.P.; Mahalingam, G.G.; Bergulst, B.; Pan, M.; Shukla, H.D.; Lasky

; Leitauer, B.; Keller, K.; Cruz, R.; Danson, M.J.; Hough, D.W.; Maddocks, D.G.; Jablo  
Jung, K.H.; Alam, M.; Freltas, T.  
Proc. Natl. Acad. Sci. U.S.A. 97, 12176-12181, 2000  
A:Authors: Hou, S.; Daniels, C.J.; Dennis, P.P.; Omer, A.D.; Ebhardt, H.; Lowe, T.M.; L  
A:Title: Genome sequence of *Halobacterium* species NRC-1.  
A:Reference number: A84160; MID:20504483  
A:Accession: D84351  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-196 <STO>  
A:Cross-references: GB:AE004437; NID:g10581437; PIDN:AAG20176.1; GSPDB:GN0138  
C:Genetics:  
A:Gene: VNG2006C  
C:Superfamily: hypothetical protein MJ0690

Query Match 68.1%; Score 32; DB 2; Length 196;  
Best Local Similarity 54.5%; Pred. No. 17;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
|||  
Db 112 SRHAQSLDR 122

RESULT 8  
F71550  
Probable endonuclease C7154 [similarity] - *Chlamydia trachomatis* (serotype D, strain UW3  
C:Species: *Chlamydia trachomatis*  
C>Date: 11-Sep-1998 #sequence\_revision 13-Sep-1998 #text\_change 04-Feb-2000  
C:Accession: F71550  
R:Stephens, R.S.; Kaiman, S.; Lammel, C.J.; Fan, J.; Marathe, R.; Aravind, L.; Mitchell,  
Science 282, 754-759, 1998  
A:Title: Genome sequence of an obligate intracellular pathogen of humans: *Chlamydia tra*  
A:Reference number: A71570; MID:99000809  
A:Accession: F71550  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-383 <ARN>  
A:Cross-references: GB:AE001289; GB:AE001273; NID:g3328550; PIDN:AAC67745.1; PID:g332855  
C:Experimental source: serotype D, strain UW-3/Cx  
C:Genetics:  
A:Gene: C7154

Query Match 68.1%; Score 32; DB 2; Length 383;  
Best Local Similarity 54.5%; Pred. No. 34;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
|||  
Db 18 SRSHRSSSR 28

RESULT 9  
S53906  
YMK1 protein - yeast (*Saccharomyces cerevisiae*)  
N:Alternate names: protein N2359; protein YNL074c  
C:Species: *Saccharomyces cerevisiae*  
C>Date: 08-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 21-Jul-2000  
C:Accession: S53906; S27290; S63006; S63935  
R:Poehlmann, R.; Philippson, P.  
submitted to the EMBL Data Library, April 1995  
A:Reference number: S53906  
A:Accession: S53906  
A:Molecule type: DNA  
A:Residues: 1-452 <POE>  
A:Cross-references: EMBL:X86470; NID:g791101; PID:g791112  
R:Taagoloff, A.  
submitted to the EMBL Data Library, January 1991  
A:Reference number: S27287  
A:Accession: S27290  
A:Molecule type: DNA

A:Residues: 1-217 <TZA>  
A:Cross-references: EMBL:X57360  
R:Poehlmann, R.; Philippson, P.  
submitted to the Protein Sequence Database, April 1996  
A:Reference number: S62997  
A:Accession: S63006  
A:Molecule type: DNA  
A:Residues: 1-452 <POW>  
A:Cross-references: EMBL:Z71350; NID:g1301960; PID:g1301961; MIPS:YNL074C  
A:Experimental source: strain S288C  
R:Poehlmann, R.; Philippson, P.  
Ieast 12, 391-402, 1996  
A:Title: Sequencing a cosmid clone of *Saccharomyces cerevisiae* chromosome XIV reveals  
A:Reference number: S63925; MID:96267764  
A:Accession: S63935  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-452 <POF>  
A:Cross-references: EMBL:X86470; NID:g791101; PID:CAA60186.1; PID:g791112  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, April 1995  
C:Genetics:  
A:Gene: SGD:YMK1  
A:Cross-references: SGD:S0005018; MIPS:YNL074C  
A:Map position: 14L

Query Match 68.1%; Score 32; DB 2; Length 452;  
Best Local Similarity 54.5%; Pred. No. 40;  
Matches 6; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
|||  
Db 189 SRSHSASATR 199

RESULT 10  
G72279  
ATP-dependent DNA helicase - *Thermotoga maritima* (strain MSB8)  
C:Species: *Thermotoga maritima*  
C>Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000  
C:Accession: G72279  
R:Neilson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MID:99287316  
A:Accession: G72279  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-648 <ARN>  
A:Cross-references: GB:AE001779; GB:AE000512; NID:g4981777; PIDN:AAD36313.1; PID:g498  
C:Experimental source: strain MSB8  
C:Genetics:  
A:Gene: TM1238  
C:Superfamily: helicase II

Query Match 68.1%; Score 32; DB 2; Length 648;  
Best Local Similarity 62.5%; Pred. No. 58;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RXHXSMEXR 9  
|||  
Db 370 RSHSLE 377

RESULT 11  
A41424  
5-methylcytosine-specific restriction enzyme A (EC 3.1.21.-) - *Escherichia coli*  
C:Species: *Escherichia coli*  
C>Date: 21-Apr-1992 #sequence\_revision 21-Apr-1992 #text\_change 04-Mar-2000  
C:Accession: A41424; S48131; D64861

R.H10m, K.; Sedgwick, S.G.  
J. Bacteriol. 173, 7368-7373, 1991  
A:Title: Cloning and structural characterization of the mcrA locus of Escherichia coli.  
A:Reference number: A41424; MUID:92041637  
A:Accession: A41424  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-277 <H10>  
A:Cross-references: GB:M76667; NID:9146793; PIDN:AAA68481.1; PID:9146794  
A:Experimental source: strain K-12  
R.Ramalingam, R.; Prasad, R.; Shivapriya, R.; Dharmalingam, K.  
J. Biosci. 17, 217-232, 1992  
A:Title: Molecular cloning and sequencing of mcrA locus and identification of McrA protein  
A:Reference number: S48131  
A:Accession: S48131  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-277 <RAM>  
A:Cross-references: EMBL:Z19104; NID:941984; PIDN:CAA79520.1; PID:941985  
R.Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: D64861  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-277 <BLAT>  
A:Cross-references: GB:AE000215; GB:U00096; NID:91787405; PIDN:AC74243.1; PID:91787406;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: mcrA; mcr; rglA  
A:Map position: 25 min  
C:Superfamily: Escherichia coli 5-methylcytosine-specific restriction enzyme A  
C:Keywords: endonuclease; hydrolase; restriction modification system

Query Match 66.0%; Score 31; DB 2; Length 277;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 RXHXSMEXR 11  
DB 69 RKHMSLDER 78

RESULT 12  
S52641  
heat shock transcription factor HSF1 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
A:Variety: Columbia  
C:Date: 01-Aug-1995 #sequence\_revision 24-Mar-1999 #text\_change 16-Jun-2000  
C:Accession: F71447; S52641; S62227; S38873  
R.Bevan, M.; Bancroft, I.; Bent, E.; Love, K.; Goodman, H.; Dean, C.; Bergkamp, R.; Dirk  
P.; Medler, H.; Medler, E.; Wambutt, R.; Weitzenegger, T.; Pohl, T.M.; Terry, N.; Giel  
vanagh, T.; Hempel, S.; Kotter, P.; Entian, K.D.; Kieger, M.; Schaeffter, M.; Funk, B.  
Nature 391, 485-488, 1998  
A:Authors: Mueller-Auer, S.; Silvey, M.; James, R.; Montfort, A.; Pons, A.; Puigdomenech  
erhoff, A.; Moores, T.; Jones, J.D.G.; Eneva, T.; Palme, K.; Benes, V.; Rechman, S.; Ans  
C.; Chalwatzis, N.  
A:Title: Analysis of 1.9 Mb of contiguous sequence from chromosome 4 of Arabidopsis thal  
A:Reference number: A71400; MUID:98121113  
A:Accession: F71447  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-495 <BEV>  
A:Cross-references: GB:Z97344; NID:92245126; PIDN:CAB10555.1; PID:92245134  
R.Huebel, A.; Schoeffl, F.  
Plant Mol. Biol. 26, 353-362, 1994  
A:Title: Arabidopsis heat shock factor: isolation and characterization of the gene and t  
A:Reference number: S52641; MUID:95036006  
A:Accession: S52641  
A:Molecule type: DNA

A:Residues: 1-215, 'M', 217-284, 'D', 286-335, 'T', 337-389, 'Y', 391-469, 474-495 <HUE>  
A:Cross-references: EMBL:X76167  
A:Accession: S62227  
A:Molecule type: mRNA  
A:Residues: 31-215, 'M', 217-284, 'D', 286-335, 'T', 337-389, 'Y', 391-469, 474-495 <HUE>  
R.Huebel, A.; Schoeffl, F.  
submitted to the EMBL Data Library, November 1993  
A:Description: Arabidopsis heat shock factor: isolation and characterization of the g  
A:Reference number: S38873  
A:Accession: S38873  
A:Molecule type: DNA  
A:Residues: 1-215, 'M', 217-284, 'D', 286-335, 'T', 337-389, 'Y', 391-461, 474-495 <HUF>  
A:Cross-references: EMBL:X76167; NID:9429154; PIDN:CAAS3761.1; PID:9429155  
C:Genetics:  
A:Gene: HSF1  
A:Map position: 4COP9-4G3845  
A:Introns: 111/3  
C:Function: <TF>  
A:Description: transcription factor that binds to heat shock promoter elements  
C:Superfamily: tomato heat shock transcription factor HSF8; HSF DNA-binding domain ho  
C:Keywords: DNA binding; heat shock; leucine zipper; nucleus; stress-induced protein;  
F.53-148/Domain: HSF DNA-binding domain homology <HSP>  
F.181-202/Region: Leucine zipper

Query Match 66.0%; Score 31; DB 1; Length 495;  
Best Local Similarity 55.6%; Pred. No. 71;  
Matches 5; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRHXHSMSE 9  
DB 277 SNSHSHSLE 285

RESULT 13  
H84075  
thioredoxin reductase BH3408 (imported) - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 03-Aug-2001  
C:Accession: H84075  
R.Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: H84075  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-330 <STO>  
A:Cross-references: GB:AP001518; GB:BA000004; NID:910175792; PIDN:BAH07127.1; GSPDB:C  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: BH3408  
C:Superfamily: thioredoxin reductase; thioredoxin reductase homology

Query Match 63.8%; Score 30; DB 2; Length 330;  
Best Local Similarity 62.5%; Pred. No. 75;  
Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RXHXHSMSE 9  
DB 190 RAHSHSVE 197

RESULT 14  
B70015  
thioredoxin reductase homolog yunc - Bacillus subtilis  
C:Species: Bacillus subtilis  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 16-Jun-2000  
C:Accession: B70015  
R.Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Ber  
C.; Bron, S.; Brouillette, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.;  
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari,

Nature 390, 249-256, 1997  
 A:Authors: Foulger, D.; Fritze, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallen  
 lech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holappel, S.; Hosono, S.; Hullo, M.F.;  
 Koelter, P.; Koningsstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,  
 A:Authors: lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel  
 Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle  
 Rieger, M.; Rivolta, C.; Roche, E.; Roche, B.; Rose, M.; Sadale, Y.; Sato, T.; Scanlon,  
 A:Authors: Schleich, S.; Schreier, R.; Scofield, F.; Sekiguchi, J.; Sekowska, A.; Seron  
 Akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama,  
 T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K  
 A:Authors: Yoshikawa, H.F.; Zunslein, E.; Yoshikawa, H.; Dancho, A.  
 A:Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.  
 A:Reference number: A69580; MUID:98044033  
 A:Accession: B70015  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-332 <KUN>  
 A:Cross-references: GB:Z99120; GB:AL009126; NID:g2635613; PIDN:CB15201.1; PID:g2635708  
 A:Experimental source: strain 168  
 C:Genetics:  
 A:Gene: yumC  
 C:Superfamily: thioresoxin reductase; thioresoxin reductase homology

Query Match 63.8%; Score 30; DB 1; Length 332;  
 Best Local Similarity 62.5%; Pred. No. 76;  
 Matches 5; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RXHXSM 9  
 | | | | |  
 Db 190 RAHEHSE 197

## RESULT 15

JC5166  
 ketol-acid reductoisomerase (EC 1.1.1.86) - *Mycobacterium avium*  
 N:Alternate names: acetoaldehyde acid isomerase  
 C:Species: *Mycobacterium avium*  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 C:Accession: JC5166  
 R:Gusberti, L.; Cantoni, R.; De Rossi, E.; Branzoni, M.; Ricciardi, G.  
 Gene 177, 83-85, 1996  
 A:Title: Cloning and sequencing of the *ilvBNC* gene cluster from *Mycobacterium avium*.  
 A:Reference number: JC5164; MUID:97080504  
 A:Accession: JC5166  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-333 <CUN>  
 A:Cross-references: GB:L49392; NID:g1196506; PIDN:AAB38428.1; PID:g1196509  
 C:Genetics:  
 A:Gene: *ilvC*  
 C:Superfamily: Methanococcus ketol-acid reductoisomerase; ketol-acid reductoisomerase hc  
 C:Keywords: Intramolecular transference; isoleucine-valine biosynthesis; isomerase; oxid  
 F;18-200/Domain: ketol-acid reductoisomerase homology <KAR>

Query Match 63.8%; Score 30; DB 1; Length 332;  
 Best Local Similarity 45.5%; Pred. No. 76;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
 | : | | | : |  
 Db 24 SOGHANSLSLR 34

Search completed: September 13, 2002, 09:24:06  
 Job time: 781 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:59 ; Search time 136.62 Seconds  
(without alignments)  
1.967 Million cell updates/sec

Title: US-09-446-543a-74  
Perfect score: 47  
Sequence: 1 SRXHXSMEXR 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

Issued\_Patents\_AA.\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCtus.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	91.5	31	3 US-09-105-678A-8	Sequence 8, Appl1
2	43	91.5	31	3 US-09-105-678A-37	Sequence 37, Appl1
3	43	91.5	31	4 US-09-172-353-4	Sequence 4, Appl1
4	43	91.5	31	4 US-08-776-971-47	Sequence 47, Appl1
5	43	91.5	31	4 US-09-421-208-8	Sequence 8, Appl1
6	43	91.5	31	4 US-09-421-208-37	Sequence 37, Appl1
7	43	91.5	32	3 US-09-105-678A-38	Sequence 38, Appl1
8	43	91.5	32	4 US-08-776-971-48	Sequence 48, Appl1
9	43	91.5	32	4 US-09-421-208-38	Sequence 38, Appl1
10	43	91.5	33	3 US-09-105-678A-39	Sequence 39, Appl1
11	43	91.5	33	4 US-08-776-971-49	Sequence 49, Appl1
12	43	91.5	33	4 US-09-421-208-39	Sequence 39, Appl1
13	43	91.5	83	4 US-08-776-971-45	Sequence 45, Appl1
14	43	91.5	83	4 US-08-776-971-124	Sequence 124, App
15	43	91.5	83	4 US-08-776-971-137	Sequence 137, App
16	42	89.4	15	4 US-08-776-971-92	Sequence 92, Appl1
17	42	89.4	29	3 US-09-105-678A-29	Sequence 29, Appl1
18	42	89.4	29	4 US-08-776-971-3	Sequence 3, Appl1
19	42	89.4	29	4 US-09-421-208-29	Sequence 29, Appl1
20	42	89.4	31	3 US-09-105-678A-7	Sequence 7, Appl1
21	42	89.4	31	3 US-08-105-678A-9	Sequence 9, Appl1
22	42	89.4	31	3 US-09-105-678A-31	Sequence 31, Appl1
23	42	89.4	31	3 US-09-105-678A-43	Sequence 43, Appl1
24	42	89.4	31	4 US-08-776-971-5	Sequence 5, Appl1
25	42	89.4	31	4 US-08-776-971-61	Sequence 61, Appl1
26	42	89.4	31	4 US-08-776-971-97	Sequence 97, Appl1
27	42	89.4	31	4 US-09-421-208-7	Sequence 7, Appl1

28	42	89.4	31	4 US-09-421-208-9	Sequence 9, Appl1
29	42	89.4	31	4 US-09-421-208-31	Sequence 31, Appl1
30	42	89.4	31	4 US-09-421-208-43	Sequence 43, Appl1
31	42	89.4	32	3 US-09-105-678A-32	Sequence 32, Appl1
32	42	89.4	32	3 US-09-105-678A-44	Sequence 44, Appl1
33	42	89.4	32	4 US-08-776-971-6	Sequence 6, Appl1
34	42	89.4	32	4 US-08-776-971-52	Sequence 52, Appl1
35	42	89.4	32	4 US-09-421-208-32	Sequence 32, Appl1
36	42	89.4	32	4 US-09-421-208-44	Sequence 44, Appl1
37	42	89.4	33	3 US-09-105-678A-33	Sequence 33, Appl1
38	42	89.4	33	3 US-09-105-678A-45	Sequence 45, Appl1
39	42	89.4	33	4 US-08-776-971-7	Sequence 7, Appl1
40	42	89.4	33	4 US-08-776-971-53	Sequence 53, Appl1
41	42	89.4	33	4 US-09-421-208-33	Sequence 33, Appl1
42	42	89.4	33	4 US-09-421-208-45	Sequence 45, Appl1
43	42	89.4	40	4 US-08-776-971-113	Sequence 113, App
44	42	89.4	87	4 US-08-776-971-59	Sequence 59, Appl1
45	42	89.4	87	4 US-08-776-971-135	Sequence 135, App

## ALIGNMENTS

```
RESULT 1
US-09-105-678A-8
: Sequence 8, Application US/09105678A
: Patent No. 6103882
:
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 1992 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
:
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/105,678A
: FILING DATE: 26-JUN-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
: US-09-105-678A-8
:
: Query Match 91.5% Score 43; DB 3; Length 31;
: Best Local Similarity 72.7%; Pred. No. 0.012;
: Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
```

OY 1 SRXHXSMEXR 11  
DB 1 SRAHQSMETR 11

## RESULT 2

US-09-105-678A-37  
; Sequence 37, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Motiya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-37

Query Match 91.5%; Score 43; DB 3; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
DB 1 SRAHQSMETR 11

## RESULT 3

US-09-172-353-4  
; Sequence 4, Application US/09172353  
; Patent No. 6197530  
; GENERAL INFORMATION:  
; APPLICANT: Stricker-Kongra, Alain  
; APPLICANT: Gu, Wei  
; TITLE OF INVENTION: GPRI0 AS A TARGET FOR IDENTIFYING WEIGHT MODULATING COMPOUNDS  
; FILE REFERENCE: 07334/102001  
; CURRENT APPLICATION NUMBER: US/09/172,353  
; CURRENT FILING DATE: 1998-10-14  
; NUMBER OF SEQ ID NOS: 7  
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 4  
; LENGTH: 31  
; TYPE: PRT  
; ORGANISM: Mus musculus  
; US-09-172-353-4

Query Match 91.5%; Score 43; DB 4; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
DB 1 SRAHQSMETR 11

## RESULT 4

US-08-776-971-47  
; Sequence 47, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Habata, Yugo  
; APPLICANT: Kawamata, Yuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Kikada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 47:

US-08-776-971-47

Query Match 91.5%; Score 43; DB 4; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
||| |||||  
Db 1 SRAHQSMETR 11

RESULT 5

US-09-421-208-8  
; Sequence 8, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-8

Query Match 91.5%; Score 43; DB 4; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
||| |||||  
Db 1 SRAHQSMETR 11

RESULT 6

US-09-421-208-37  
; Sequence 37, Application US/09421208  
; Patent No. 6258561

GENERAL INFORMATION:

; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-37

Query Match 91.5%; Score 43; DB 4; Length 31;  
Best Local Similarity 72.7%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
||| |||||  
Db 1 SRAHQSMETR 11

RESULT 7

US-09-105-678A-38  
; Sequence 38, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-38

Query Match 91.5% Score 43; DB 3; Length 32;  
Best Local Similarity 72.7% Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
||| |||||  
DB 1 SRAHQSMETR 11

RESULT 8  
US-08-776-971-48  
Sequence 48, Application US/08776971B  
Patent No. 6228964  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 48:  
US-08-776-971-48

Query Match 91.5% Score 43; DB 4; Length 32;  
Best Local Similarity 72.7% Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
||| |||||  
DB 1 SRAHQSMETR 11

RESULT 9  
US-09-421-208-38  
Sequence 38, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid

STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-38

Query Match 91.5%; Score 43; DB 4; Length 32;  
Best Local Similarity 72.7%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
Db 1 SRAHQSMETR 11

RESULT 10  
US-09-105-678A-39  
; Sequence 39, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-39

Query Match 91.5%; Score 43; DB 3; Length 33;  
Best Local Similarity 72.7%; Pred. No. 0.013;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
Db 1 SRAHQSMETR 11

RESULT 11  
US-08-776-971-49  
; Sequence 49, Application US/08776971B

Patent No. 6228984  
; GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
APPLICANT: Habata, Yugo  
APPLICANT: Kawamata, Yuji  
APPLICANT: Hosoya, Masaki  
APPLICANT: Fujii, Ryo  
APPLICANT: Fukusumi, Shoji  
APPLICANT: Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 49:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 49:  
US-08-776-971-49

Query Match 91.5%; Score 43; DB 4; Length 33;  
Best Local Similarity 72.7%; Pred. No. 0.013;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 SRXHXSMEXR 11  
Db 1 SRAHQSMETR 11

RESULT 12  
US-09-421-208-39  
; Sequence 39, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 39:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-39

Query Match 91.5%; Score 43; DB 4; Length 33;  
Best Local Similarity 72.7%; Pred. No. 0.013;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
DB 1 SRAHSHMETR 11

RESULT 13  
US-08-776-971-45  
Sequence 45, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette

COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 45:  
US-08-776-971-45

Query Match 91.5%; Score 43; DB 4; Length 83;  
Best Local Similarity 72.7%; Pred. No. 0.032;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
DB 22 SRAHSHMETR 32

RESULT 14  
US-08-776-971-124  
Sequence 124, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 124:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 124:  
US-08-776-971-124

Query Match 91.5%; Score 43; DB 4; Length 83;  
Best Local Similarity 72.7%; Pred. No. 0.032;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
||| |||||  
DB 22 SRAHQSMETR 32

RESULT 15  
US-08-776-971-137  
Sequence 137, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 137:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 83 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 137:  
US-08-776-971-137

Query Match 91.5%; Score 43; DB 4; Length 83;  
Best Local Similarity 72.7%; Pred. No. 0.032;  
Matches 8; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 SRXHXSMEXR 11  
||| |||||  
DB 22 SRAHQSMETR 32

Search completed: September 13, 2002, 09:20:59  
Job time: 624 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:09:25 ; Search time 25.73 seconds  
(without alignments)  
221.875 Million cell updates/sec

Title: SEQ74-FUSED-TO-SEQ73  
Perfect score: 156  
Sequence: 1 SRXHHSMEXRTPDINPAWXXRGIRPYGRFX 33

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :  
1: SP archaea:\*  
2: SP bacteria:\*  
3: SP fungi:\*  
4: SP human:\*  
5: SP invertebrate:\*  
6: SP mammal:\*  
7: SP mhc:\*  
8: SP organelle:\*  
9: SP phage:\*  
10: SP plant:\*  
11: SP rodent:\*  
12: SP virus:\*  
13: SP vertebrate:\*  
14: SP unclassified:\*  
15: SP virus:\*  
16: SP bacteriophage:\*  
17: SP archaea:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Length	DB	ID	Description
1	98	62.8	117	13	Q9W624	Q9W624 carassius a
2	50	32.1	692	2	Q9L8J6	Q9L8J6 rhodospirill
3	49.5	31.4	664	16	Q9L1W4	Q9L1W4 pseudomonas
4	49	31.4	220	16	Q9I022	Q9I022 pseudomonas
5	49	30.1	790	10	Q9M371	Q9M371 arabidopsis
6	47	30.1	128	16	P74747	P74747 synecocyst
7	47	30.1	173	17	Q97A00	Q97A00 thermoplasma
8	47	30.1	1271	2	Q93SV4	Q93SV4 chlorobium
9	46.5	29.8	420	2	Q93LZ7	Q93LZ7 streptomyces
10	46	29.5	54	4	Q9U0F9	Q9U0F9 homo sapien
11	46	29.5	277	11	Q9D8A2	Q9D8A2 mus musculu
12	46	29.5	286	11	Q9DAE5	Q9DAE5 mus musculu
13	46	29.5	314	16	O06348	O06348 mycobacteri
14	46	29.5	333	16	Q9PH76	Q9PH76 xyella fas
15	46	29.5	348	5	O45431	O45431 caenorhabdi
16	46	29.5	450	11	Q9DA19	Q9DA19 mus musculu

17	46	29.5	465	4	O60687	O60687 homo sapien
18	46	29.5	540	10	Q9L6Z0	Q9L6Z0 oryza sativ
19	46	29.5	798	11	Q9LW50	Q9LW50 mus musculu
20	46	29.5	1236	2	Q9JPR4	Q9JPR4 rhodocyclu
21	46	29.5	1292	2	Q9ZGE5	Q9ZGE5 heliobacill
22	45	28.8	240	16	Q9RSN5	Q9RSN5 delinococcus
23	45	28.8	412	10	Q9FYNO	Q9FYNO arabidopsis
24	45	28.8	501	10	Q9LYX0	Q9LYX0 arabidopsis
25	45	28.8	637	16	Q9A5E9	Q9A5E9 caulobacter
26	45	28.8	767	5	Q20170	Q20170 caenorhabdi
27	44.5	28.5	954	16	Q9KGP0	Q9KGP0 bacillus ha
28	44	28.2	72	2	Q9EYB3	Q9EYB3 escherichia
29	44	28.2	145	5	Q95ZB5	Q95ZB5 leishmania
30	44	28.2	159	16	Q9PD41	Q9PD41 xyella fas
31	44	28.2	250	16	Q9TDF2	Q9TDF2 pseudomonas
32	44	28.2	284	17	O50128	O50128 pyrococcus
33	44	28.2	419	4	Q9Y276	Q9Y276 homo sapien
34	44	28.2	518	10	Q9LVE2	Q9LVE2 arabidopsis
35	44	28.2	548	10	Q9LFA0	Q9LFA0 arabidopsis
36	44	28.2	762	5	Q9V748	Q9V748 drosophila
37	44	28.2	881	5	Q9NE42	Q9NE42 leishmania
38	43.5	27.9	503	16	Q9KRY1	Q9KRY1 vibrio chol
39	43.5	27.9	852	5	O01978	O01978 caenorhabdi
40	43.5	27.9	1501	10	Q9SD86	Q9SD86 arabidopsis
41	43	27.6	228	17	Q9HJT2	Q9HJT2 thermoplasma
42	43	27.6	232	16	Q9R221	Q9R221 dermoplasma
43	43	27.6	309	5	O17234	O17234 caenorhabdi
44	43	27.6	327	10	Q94K09	Q94K09 brassica ca
45	43	27.6	327	10	Q94FZ9	Q94FZ9 brassica ca

## ALIGNMENTS

RESULT	ID	Query Match	Score	DB	Length
Q9W624	Q9W624	62.8%	98	13	117
AC	Q9W624	Best Local Similarity	53.8%	Pred. No. 3e-08	
DT	01-NOV-1999 (Trembl)	Matches 14; Conservative	8; Mismatches	4; Indels	0; Gaps
DT	01-DEC-2001 (Trembl)				
DE	C-RE AMIDE.				
OS	Carassius auratus (Goldfish).				
OC	Eukaryota: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi:				
OC	Actinopterygii: Neopterygii: Teleostei: Euteleostei: Ostariophysi:				
OC	Cypriniformes: Cyprinidae: Carassius.				
OX	NCBI_TaxID=7957;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-BRAIN:				
RA	Satake H., Minakata H., Fujimoto M.;				
RT	"Carassius Rhamde (C-RE amide).";				
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AB020024; BAA76662.1;				
SQ	SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;				

Query Match	Score	DB	Length
Best Local Similarity	53.8%	Pred. No. 3e-08	
Matches 14; Conservative	8; Mismatches	4; Indels	0; Gaps
OY	6 HSMXRTDINPAWXXRGIRPYGRF 31		
DB	50 HNDNRSPETDPYVGRGVPIGRF 75		
RESULT	2		
Q9L8J6	Q9L8J6	PRELIMINARY;	PRT: 692 AA.
AC	Q9L8J6		
DT	01-OCT-2000 (Trembl)		
DT	01-OCT-2000 (Trembl)		
DT	01-OCT-2000 (Trembl)		
DT	01-JUN-2001 (Trembl)		

	Query Match	Similarity	Score	DB	Length
Db	Best local	10	62.58	Pred. No. 1	664
	Matches	10	Conservative	1	Mismatches 4: Indels 1: Gaps 1
Qy	12	TPDINPAMYXXRGIRP	27		1 1 1 1 1 1
	478	TPDINP-WFIQRRSGRP	492		

Query Match	31.48;	Score 49;	DB 10;	Length 790;
Best Local Similarity	47.48;	Pred. No. 20;	.	

	Query Match	30.1%	Score 47:	DB 16;	Length 128;
	Best Local Similarity	33.3%;	Pred. NO. 5.9;		
Matches	11; Conservative	5;	Mismatches	13;	Indels 4;
	Gaps	1;			
Oy	1 SRXHXHSEKRTPDINPAMYXXR---GIRPYG	29			
	:::				
bj	11 SPENNOSLDCSSPPPSRKMYQARLEFSFIRTPG	43			

RESULT	7			
Q97A00				
ID	Q97A00	PRELIMINARY;	PRT;	173 AA.
AC	Q97A00;			
DT	01-OCT-2001 (TReMBLrel. 18, Created)			
DT	01-OCT-2001 (TReMBLrel. 18, last sequence update)			
DT	01-DEC-2001 (TReMBLrel. 19, last annotation update)			
DE	FERRIDYOCHELIN BINDING PROTEIN.			
GN	TW61032219.			
OS	Thermoplasma volcanium.			
OC	Archaea; Euryarchaeota; Thermoplasmales; Thermoplasmataceae;			
OC	Thermoplasma			
OX	NCBI_TaxID=50339;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=GS51 / DSM 4299 / JCM 9571;			
RX	MEDLINE=20570466; PubMed=11121031;			
RA	Kawashima T., Anano N., Koike H., Makino S.-I., Higuchi S.,			
RA	Kawashima-Ohya Y., Watanabe K., Yamazaki M., Kanehori K., Kawamoto T.,			
RA	Nunoshiba T., Yamamoto Y., Arimaki H., Makino K., Suzuki M.;			
RT	"Archaeal adaptation to higher temperatures revealed by genomic			
RT	sequence of Thermoplasma volcanium."			
RL	Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).			
DR	EMBL; AP000995; BAB60152.1; -			
DR	InterPro: IPR001451; Hexapep_transf.			
DR	Pfam: PF00132; hexapep; 4.			
DR	Complete proteome.			
QO	SEQUENCE 173 AA; 19041 MW; 1A750DDB821C7524 CRC64;			

RESULT	ID	PRELIMINARY:	PRT:	1271 AA.
093SV4	093SV4			
AC	093SV4			
DT	01-DEC-2001	(TREMBLrel. 19, Created)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)		
DE	BCHH1.			
GN	BCHH1.			
OS	Chlorobium tepidum.			
OC	Bacteria; Green sulfur bacteria; Chlorobium.			
OX	NCBI_TaxID:1097;			
RN	[1]			
RP	SEQUENCE FROM N.A..			
RX	MEDLINE=20433268; PubMed=10976061;			
RA	Xiong J., Fischer W.M., Inoue K., Nakahara M., Bauer C.E.;			
RT	"Molecular evidence for the early evolution of photosynthesis."			
RL	Science 289:1724-1730(2000).			
SE	EMBL; AY005136; AAC12412.1; -			
SO	SEQUENCE 1271 AA; 141905 MW; 752C9D216549A2B1 CRC64;			

Query Match	30.1%	Score 47	DB 2	Length 1271
Best Local Similarity	32.1%	Pred. No. 69		
Matches	9	Conservative	5	Mismatches
			8	Indels
				Gaps
				1
Oy	7	SMEKRTPDINPAMWXX-----RGIRPV	28	
		:::     ::	1-1	
Db	1152	ALERTKLLNPWYEMATEHGEGYREI	1179	

RESULT	9		
093JLZ7			
ID	093JLZ7	PRELIMINARY;	PRF: 420 AA.
AC	093JLZ7;		
DT	01-DEC-2001	(TREMBLrel. 19, Created)	
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)	
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	
DE	CHAIN LENGTH FACTOR-LIKE PROTEIN.		
GN	AUR2B.		
OS	Streptomyces aureofaciens.		
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;		
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.		
OX	NCBI_TaxID=1894;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CCM3239;		
RA	Kormaneec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;		
RT	"Cloning and characterization of a new polyketide gene cluster in		
RT	Streptomyces aureofaciens CCM3239."		
RL	Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.		
EMBL	AY033994; AAK61719.1;		
SD	SEQUENCE 420 AA; 43011 MW; 3627822BE86C2DEA CRC64;		

Query Match	29.8%	Score 46.5	DB 2	Length 420
Best Local Similarity	52.6%	Pred. No. 25		
Matches 10	Conservative 1	Mismatches 3	Indels 5	Gaps 1
QY	18 AWYXX-----RCIRPVGRR	31		
db	37 AWMAAVLRCGSGIRPVGRR	55		

```

RESULT 10
ID 09UJF9 PRELIMINARY; PRT: 54 AA.
AC 09UJF9:
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DE 01-MAY-2000 (TREMBLrel. 13, Last annotation update)
DE DJ479J7.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).
GN DJ479J7.3.
OS Homo sapiens (Human)
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN 11
RP SEQUENCE FROM N.A.
RA Lawlor S.;
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL035608; CAB55682.1; -.
FT NON_TER 54
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 29.5%; Score 46; DB 4; Length 54;
Best Local Similarity 43.8%; Pred. No. 3.4;
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 12 TPDINPAMYXXRGIRP 27
Db 18 TPVPTPTWYAGSGYYP 33

RESULT 11
ID 09DA2 PRELIMINARY; PRT: 277 AA.
AC 09DA2:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 1700023B02RIK PROTEIN.
GN 1700023B02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Wyszynaw-Borls A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK008242; BAB2553.1; -.
MGD: MGI:1914185; 1700023B02RIK.
SQ SEQUENCE 277 AA; 31797 MW; 4BEAD95516FBAC32 CRC64;

```

```

Query Match 29.5%; Score 46; DB 11; Length 277;
Best Local Similarity 43.3%; Pred. No. 20;
Matches 13; Conservative 1; Mismatches 10; Indels 6; Gaps 1;

OY 1 SRXHXSMEXRPDPINPAMYXXRGIRPVG 30
Db 200 SRSHRSPEKKSDDRN-----RGIRSR 223

RESULT 12
ID 09DA5 PRELIMINARY; PRT: 286 AA.
AC 09DA5:
DT 01-JUN-2001 (TREMBLrel. 17, Created)
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)
DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)
DE 1700023B02RIK PROTEIN.
GN 1700023B02RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN 11
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=TESTIS;
RX MEDLINE=21085660; PubMed=11217851;
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
RA Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K. I.,
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
RA Fleschmann W., Gaasterland T., Glass C., King B., Kochiwa H.,
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,
RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,
RA Sakai K., Okido T., Furuno M., Aono H., Balderelli R., Barsh G.,
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamlya M., Lee N.H.,
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,
RA Suzuki H., Toyo-oka K., Wang K.H., Weltz C., Whitaker C., Wilming L.,
RA Wyszynaw-Borls A., Yoshida K., Hasegawa Y., Kawai H., Kohsaki S.,
RA Hayashizaki Y.;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
DR EMBL; AK005900; BAB24307.1; -.
MGD: MGI:1914185; 1700023B02RIK.
SQ SEQUENCE 286 AA; 32818 MW; 068003C5E894827B CRC64;

Query Match 29.5%; Score 46; DB 11; Length 286;
Best Local Similarity 43.3%; Pred. No. 20;
Matches 13; Conservative 1; Mismatches 10; Indels 6; Gaps 1;

OY 1 SRXHXSMEXRPDPINPAMYXXRGIRPVG 30
Db 209 SRSHRSPEKKSDDRN-----RGIRSR 222

RESULT 13
ID 006348 PRELIMINARY; PRT: 314 AA.
AC 006348:
DT 01-JUL-1997 (TREMBLrel. 04, Created)
DT 01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHETICAL 33.2 KDA PROTEIN.
GN RV3485C OR MTCY13E12.38C.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

```

OC Actinomycetales; Corynebacteriineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE=96895987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eigmler K., Gas S., Barry C.E. III, Tekala F.,  
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
 Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,  
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 Rutter S., Seeger K., Skellon S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 complete genome sequence."  
 RL Nature 393:537-544(1998).  
 CC -1 SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
 CC (SDR) FAMILY.  
 CC EMBL: 295390; CAB08708.1; -.  
 DR HSSP: P29132; IDFT.  
 DR Tuberculist: RV3485c; -.  
 DR InterPro: IPR002198; ADH\_short.  
 DR Pfam: PF00106; adh\_short; 1.  
 DR PROSITE: PS00061; ADH\_SHORT; UNKNOWN\_1.  
 KW Complete proteome; Hypothetical protein; Oxidoreductase.  
 SO SEQUENCE 314 AA; 33194 MW; 2614BA917E09274 CRC64;

Query Match 29.5%; Score 46; DB 16; Length 314;  
 Best Local Similarity 31.8%; Pred. No. 22;  
 Matches 7; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 6 HSMEXRTPDINPAMXXRCIRP 27  
 I : : : : :  
 Db 205 HMKLADELGPSWVRVNSIRP 226

RESULT 14  
 O9PH76 PRELIMINARY; PRT; 333 AA.  
 AC O9PH76;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYDROXYBENZONATE OCTAPRENYLTRANSFERASE.  
 CN XF0068.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OX NCBI\_TaxID=2371;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Rehnach F.C., Artuda P., Abreu F.A., Acencio M.,  
 Alvarenga R., Alves L.M.C., Araya J.E., Baia J.E., Battista C.S.,  
 Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bionnes M.R.S.,  
 Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
 Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.H.,  
 Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 Pachincun A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,  
 Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 Ho P.L., Hohensei J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
 Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vitorello C.B.,  
 Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A.Jr., Nodrega F.G., Nunes L.R., Oliveira M.A.,  
 de Oliveira M.C., de Oliveira R.C., Palmeri D.A., Paris A.,

RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silva J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa";  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003860; AA62881.1; -.  
 DR InterPro: IPR000537; UBIA.  
 DR Pfam: PF01040; UBIA; 1.  
 DR PROSITE: PS00943; UBIA; UNKNOWN\_1.  
 KW Complete proteome.  
 SO SEQUENCE 333 AA; 37931 MW; ECF3FA716C962B95 CRC64;

Query Match 29.5%; Score 46; DB 16; Length 333;  
 Best Local Similarity 53.3%; Pred. No. 24;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 15 INPAMXXRCIRPVG 29  
 : : : : :  
 Db 54 LDPYWKLAGDRPVG 68

RESULT 15  
 ID 045431 PRELIMINARY; PRT; 348 AA.  
 AC 045431;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE F32B6.3 PROTEIN.  
 DE GN F32B6.3.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditidae; Rhabditidae;  
 OC Rhabditidae; Pelodierinae; Caenorhabditis.  
 OX NCBI\_TaxID=6239;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Basham V.M.;  
 RL Submitted (OCT-1996) to the EMBL/Genbank/DBJ databases.  
 RN 121  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9651916;  
 RT none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL: 281074; CAB03039.1; -.  
 DR InterPro: IPR004098; PfP18.  
 DR InterPro: IPR003648; SFM.  
 DR Pfam: PF02840; PfP18; 1.  
 DR SMART: SM00500; SFM; 1.  
 SO SEQUENCE 348 AA; 40152 MW; 8EFCF68312BB401 CRC64;

Query Match 29.5%; Score 46; DB 5; Length 348;  
 Best Local Similarity 38.9%; Pred. No. 25;  
 Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 2 RXHXSMEXRTPDINPAM 19  
 I : : : : :  
 Db 112 RRRLHQLLEADPLNEGW 129

Search completed: September 13, 2002, 09:11:32  
 Job time: 127 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:10:10 : Search time 10.33 Seconds  
(without alignments)  
123.693 Million cell updates/sec

Title: SEQ74-FUSED-TO-SEQ73  
Perfect score: 156  
Sequence: 1 SRXHXHSMEXRTPDINPAMYXXRGIRPVCGRFX 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	147	94.2	83	PRRP_RAT
2	147	94.2	87	PRRP_HUMAN
3	146	93.6	98	PRRP_BOVIN
4	46	29.5	676	EXLI_HUMAN
5	46	29.5	798	UNR_RAT
6	45	28.8	428	NER3_BOVIN
7	44.5	28.5	240	PLSC_HELPI
8	44	28.2	288	Y587_PASMU
9	44	28.2	428	NER3_HUMAN
10	44	28.2	962	UVRA_METMO
11	43	27.6	376	OPSL_LIMPO
12	43	27.6	376	OPSL_LIMPO
13	43	27.6	383	CYCR_CHRYI
14	43	27.6	476	YAAU_ECOLI
15	43	27.6	581	POL_MLVAK
16	43	27.6	719	NRPI_YEAST
17	43	27.6	843	POL_MLVAK
18	43	27.6	1196	POL_MLVAK
19	43	27.6	1196	POL_MLVAK
20	42	26.9	402	EX7L_STRCO
21	42	26.9	926	UVRA_AOUAE
22	42	26.9	986	CYGR_ARBP
23	41.5	26.6	345	ARGC_BACHD
24	41.5	26.6	1882	Y468_MYCPN
25	41	26.3	162	PHAI_FREDI
26	41	26.3	162	PHAI_PSEAI
27	41	26.3	162	PHAI_PSEAI
28	41	26.3	162	PHAI_PSEAI
29	41	26.3	162	PHAI_PSEAI
30	41	26.3	162	PHAI_PSEAI
31	41	26.3	162	PHAI_PSEAI
32	41	26.3	162	PHAI_PSEAI
33	41	26.3	162	PHAI_PSEAI

34	41	26.3	441	1	HYDG_ECOLI	P14375	escherichia
35	41	26.3	445	1	CLUS_CANPA	P25473	canis famil
36	41	26.3	510	1	YCGE_ECOLI	P29013	escherichia
37	41	26.3	798	1	UNR_HUMAN	O75534	homo sapien
38	41	26.3	973	1	UVRA_RHIME	P56889	rhizobium m
39	41	26.3	1068	1	FAK1_XENLA	O91738	xenopus lae
40	41	26.3	1083	1	T2D3_HUMAN	O00268	homo sapien
41	40.5	26.0	258	1	BOX5_NOTVI	P53771	notophthalm
42	40.5	26.0	374	1	YHHI_ECOLI	P31993	escherichia
43	40.5	26.0	860	1	WG12_BPR03	O37893	bacterioph
44	40	25.6	273	1	TC1A_CAEBL	P03934	caenorhabdi
45	40	25.6	329	1	Y493_MYCTU	O11158	mycobacteri

## ALIGNMENTS

RESULT 1	PRRP_RAT	STANDARD:	PRT:	83 AA.
AC	P81278			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP20].			
DE	releasing peptide PRRP20].			
GN	PRH.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCBI_Taxid:10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	MEDLINE=98268781; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,			
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,			
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RT	"A prolactin-releasing peptide in the brain.";			
RL	Nature 393:272-276(1996).			
RN	[2]			
RP	TISSUE SPECIFICITY.			
RP	PubMed=10498338;			
RX	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Hinuma S.,			
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,			
RA	Sumino Y., Fujino M.;			
RT	"Tissue distribution of prolactin-releasing peptide (PRRP) and its			
RT	receptor.";			
RL	Regul. Pept. 83:1-10(1999).			
CC	-1- FUNCTION: Stimulates prolactin (PRL) release and regulates the			
CC	expression of prolactin through its receptor GPR10. May stimulate			
CC	lactotrophs directly to secrete PRL.			
CC	-1- TISSUE SPECIFICITY: Widely expressed, with highest levels in			
CC	medulla oblongata and hypothalamus.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).			
CC	-----			
DR	EMBL; AB015418; BA29026.1;			
KW	Hormone; Amidation; Signal;			
FT	SIGNAL			
FT	PEPTIDE			
FT	PEPTIDE			
FT	MOD_RES			
FT	SEQUENCE			
SO	83 AA: 9215 MW: DQC75A264EBB4F29 CRC64;			

Query Match 94.2%; Score 147; DB 1; Length 83;  
 Best Local Similarity 83.9%; Pred. No. 1.3e-16;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31  
 DB 22 SRAHQSHMETRTPDINPAWYTGIRPVGRF 52

## RESULT 2

PRRP\_HUMAN STANDARD: PRT: 87 AA.

AC 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-releasing peptide PRRP20].  
 GN PRR.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RM [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujil R., Kawamata Y., Hosoya M., Fukusumi S.,

RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,

RA Kurokawa T., Nishimura O., Onda H., Fujino M.,

RA "A proactin-releasing peptide in the brain.";

RT Nature 393:272-276(1998).

RL [2]

RM TISSUE SPECIFICITY.

RP PubMed=10498338;

RA Fujil R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,

RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,

RA "Tissue distribution of proactin-releasing peptide (PRRP) and its

RT receptor.";

RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates proactin (PRL) release and regulates the

CC expression of proactin through its receptor GPR10. May stimulate

CC lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AB015419; BAA29027.1; -.

DR MIM: 602663; -.

KW Hormone; Amidation; Signal.

FT SIGNAL 1 22 BY SIMILARITY.

FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP31.

FT MOD\_RES 34 53 PROACTIN-RELEASING PEPTIDE PRRP20.

FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).

SO SEQUENCE 87 AA; 9639 MW; 229A2F350CF981B CRC64;

Query Match 94.2%; Score 147; DB 1; Length 87;  
 Best Local Similarity 83.9%; Pred. No. 1.3e-16;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31  
 DB 23 SRAHQSHMETRTPDINPAWYASRGIRPVGRF 53

RESULT 3  
 PRRP\_BOVIN STANDARD: PRT: 98 AA.

AC P81264;

DT 30-MAY-2000 (Rel. 39, Created)

DT 30-MAY-2000 (Rel. 39, Last sequence update)

DT 01-MAR-2002 (Rel. 41, Last annotation update)

DE Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-releasing peptide PRRP20].

GN PRR.

OS Bos taurus (bovine).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RM [1]

RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RC TISSUE=Brain;

RA MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujil R., Kawamata Y., Hosoya M., Fukusumi S.,

RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,

RA Kurokawa T., Nishimura O., Onda H., Fujino M.,

RA "A proactin-releasing peptide in the brain.";

RT Nature 393:272-276(1998).

RL -1- FUNCTION: Stimulates proactin (PRL) release and regulates the

CC expression of proactin through its receptor GPR10. May stimulate

CC lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

DR EMBL: AB015417; BAA29025.1; -.

KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT SIGNAL 1 22

FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP31.

FT MOD\_RES 33 53 PROACTIN-RELEASING PEPTIDE PRRP20.

FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).

SO SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 93.6%; Score 146; DB 1; Length 98;  
 Best Local Similarity 83.9%; Pred. No. 2.2e-16;  
 Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31  
 DB 23 SRAHQSHMETRTPDINPAWYAGRGIRPVGRF 53

## RESULT 4

EX1L\_HUMAN STANDARD: PRT: 676 AA.

AC Q92935;

DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Exostosis-like 1 (Exostosis-L) (Multiple exostosis-like protein).

GN EXTL1 OR EXTL.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;

RM [1]

RP SEQUENCE FROM N.A.

RA MEDLINE=97189339; PubMed=9037597;



RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
RT "Identification and localization of the gene for EXT1, a third member  
RT of the multiple exostoses gene family.",  
RL Genome Res. 7:10-16(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Xu L., Deng H.X., Xia J.H., Pan O., Liu C.Y.;  
RT "Mutations of the EXT genes in hereditary multiple exostoses in  
RT Chinese";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Wuyts W., Spieker N., Van Roy N., De Paeye A., De Bouille K.,  
RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;  
RT "Refined physical mapping and genomic structure of the EXT1 gene";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
CC reticulum (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: U67191; AAC51141.1; -.  
DR EMBL: AF083633; AAD02840.1; -.  
DR EMBL: AF083623; AAD02840.1; JOINED.  
DR EMBL: AF083624; AAD02840.1; JOINED.  
DR EMBL: AF083625; AAD02840.1; JOINED.  
DR EMBL: AF083626; AAD02840.1; JOINED.  
DR EMBL: AF083627; AAD02840.1; JOINED.  
DR EMBL: AF083628; AAD02840.1; JOINED.  
DR EMBL: AF083629; AAD02840.1; JOINED.  
DR EMBL: AF083630; AAD02840.1; JOINED.  
DR EMBL: AF083631; AAD02840.1; JOINED.  
DR EMBL: AF083632; AAD02840.1; JOINED.  
DR EMBL: AF153960; AAF73172.1; -.  
DR EMBL: AF151391; AAF73172.1; JOINED.  
DR MIM: 601738; -.  
DR InterPro: IPR004263; Exostosin.  
DR Pfam: PF03016; Exostosin; 1.  
KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.  
FT TRANSMEM 10 30  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
FT CARBOHYD 269 269  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 676 AA; 74673 MW; B5E006A8762E5633 CRC64;  
  
Query Match 29.5%; Score 46; DB 1; Length 676;  
Best Local Similarity 45.0%; Pred. No. 11;  
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
  
Oy 12 TPDIWPAMYXXRGIRPVGRF 31  
Db 400 SPQDFPFYFYLQGSRRPGRF 419  
  
RESULT 5  
UNR\_RAT  
ID UNR\_RAT STANDARD: PRT: 798 AA.  
AC P18395;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE UNR protein.  
GN UNR.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE=90370473; PubMed=2204029;  
RA Jeffers M., Pacilucci R., Pellicer A.;  
RT "Characterization of unr: a gene closely linked to N-ras";  
RL Nucleic Acids Res. 18:4891-4899(1990).  
CC -1- FUNCTION: RNA-BINDING PROTEIN (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE COLD-SHOCK DOMAIN (CSD) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: X52311; CA36549.1; -.  
DR PIR: S11210; S11210.  
DR HSSP: P15277; IMJC.  
DR InterPro: IPR002059; Cold\_shock.  
DR Pfam: PF00313; CSD; 8.  
DR SMART: SM00357; CSD; 5.  
DR PROSITE: PS00352; COLD\_SHOCK; 4.  
KW RNA-binding; Repeat.  
FT DOMAIN 26 87 CSD 1.  
FT DOMAIN 136 179 CSD 2 (INCOMPLETE).  
FT DOMAIN 186 245 CSD 3.  
FT DOMAIN 297 337 CSD 4 (INCOMPLETE).  
FT DOMAIN 349 410 CSD 5.  
FT DOMAIN 447 507 CSD 6.  
FT DOMAIN 519 579 CSD 7.  
FT DOMAIN 610 670 CSD 8.  
FT DOMAIN 674 735 CSD 9.  
SQ SEQUENCE 798 AA; 88894 MW; FA483FAB809595A4 CRC64;  
  
Query Match 29.5%; Score 46; DB 1; Length 798;  
Best Local Similarity 39.1%; Pred. No. 13;  
Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;  
  
Oy 6 HSMEXRTPDINPAMYXXRGIRPV 28  
Db 583 HSVNGITEANPTIYSGKVRPL 605  
  
RESULT 6  
NER3\_BOVIN  
ID NER3\_BOVIN STANDARD: PRT: 428 AA.  
AC O97859;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Sialidase 3 (EC 3.2.1.18) (Membrane sialidase) (ganglioside sialidase)  
DE (N-acetyl-alpha-neuraminidase 3).  
GN NEU3.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=99143165; PubMed=9988745;  
RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,  
RA Savada M.;  
RT "Molecular cloning and characterization of a plasma membrane-associated sialidase specific for gangliosides.";

```

RL J. Biol. Chem. 274:5004-5011(1999).
CC -1- FUNCTION: Plays a role in modulating the ganglioside content of
CC the lipid bilayer at the level of membrane-bound sialyl
CC glycoconjugates.
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and
CC synthetic substrates.
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).
CC -1- TISSUE SPECIFICITY: Expressed in brain.
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.
CC -1- SIMILARITY: CONTAINS 3 BNR REPEATS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: AB008184; BAA75071.1; -
DR InterPro: IPR002860; BNR.
DR Pfam: PF02012; BNR.
KW Hydrolase; Glycosidase; Membrane; Repeat.
FT REPEAT 129 140 BNR 1.
FT REPEAT 203 214 BNR 2.
FT REPEAT 254 265 BNR 3.
FT SITE 24 27 PRIP MOTIF.
FT ACT_SITE 25 25 BY SIMILARITY.
FT ACT_SITE 45 45 POTENTIAL.
FT ACT_SITE 50 50 POTENTIAL.
FT ACT_SITE 87 87 POTENTIAL.
FT ACT_SITE 225 225 POTENTIAL.
FT ACT_SITE 245 245 POTENTIAL.
FT ACT_SITE 341 341 BY SIMILARITY.
FT ACT_SITE 371 371 POTENTIAL.
FT ACT_SITE 388 388 POTENTIAL.
SQ SEQUENCE 428 AA; 47916 MW; 418B34F3245A8F21 CRC64;

Query Match 28.8%; Score 45; DB 1; Length 428;
Best Local Similarity 29.6%; Pred. No. 10;
Matches 8; Conservative 6; Mismatches 13; Indels 0; Gaps 0;

QY 2 RKXHSMEKXRTPDINPAWYXXRGIRPV 28
DB 195 RARPHSLMITYSDLGATWGHGRILKPM 221

RESULT 7
ID PLSC_HELPHY STANDARD; PRT; 240 AA.
AC 025903:
DT 15-DEC-1998 (Rel. 37, Created)
DT 15-DEC-1998 (Rel. 37, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 1-acetyl-sn-glycerol-3-phosphate acyltransferase (EC 2.3.1.51) (1-AGP
DE acyltransferase) (1-AGPAT) (Lysophosphatidic acid acyltransferase)
DE (LIPAT)
GN PLSC OR HP1348.
OS Helicobacter pylori (Campylobacter pylori).
OC Bacteria; Proteobacteria; epsilon subdivision; Helicobacter group;
OC Helicobacter.
OX NCBI_TaxID=210;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=26695 / ATCC 700392;
RX MEDLINE=9794467; PubMed=9252185;
RA Tomb J.-F., White O., Kerlavage A.R., Clayton R.A., Sutton G.G.,
RA Fleischmann R.D., Ketchum K.A., Klein H.-P., Gill S., Dougherty B.A.,
RA Nelson K., Quackenbush J., Zhou L., Kirkness E.F., Peterson S.,
RA Loftus B., Richardson D., Dodson R., Khairak H.G., Giodex A.,

```

```

RA McKenney K., Fitzgerald L.M., Lee N., Adams M.D., Hickey E.K.,
RA Berg D.E., Gocayne J.D., Utterback T.R., Peterson J.D., Kelley J.M.,
RA Cotton M.D., Weldman J.M., Fujii C., Bowman C., Walley L., Wallin E.,
RA Hayes W.S., Borodovsky M., Karp P.D., Smith H.O., Fraser C.M.,
RA Venter J.C.;
RT "The complete genome sequence of the gastric pathogen Helicobacter
RT pylori."
RL Nature 388:539-547(1997).
CC -1- FUNCTION: CONVERTS LYSOPHOSPHATIDIC ACID (LPA) INTO PHOSPHATIDIC
CC ACID BY INCORPORATING ACYL MOIETY AT THE 2 POSITION.
CC -1- CATALYTIC ACTIVITY: Acyl-CoA + 1-acetyl-sn-glycerol 3-phosphate -
CC CoA + 1,2-diacyl-sn-glycerol 3-phosphate.
CC -1- PATHWAY: SECOND STEP IN DE NOVO PHOSPHOLIPID BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Inner membrane-associated (Potential).
CC -1- SIMILARITY: BELONGS TO THE 1-ACYL-SN-GLYCEROL-3-PHOSPHATE
CC ACYTRANSFERASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----
DR EMBL: AE000636; AAD08393.1; -
DR TIGR: HP1348; -
DR InterPro: IPR002123; Acyltransferase.
DR Pfam: PF01553; Acyltransferase; 1.
KW Phospholipid biosynthesis; Transferase; Acyltransferase;
KW Inner membrane; Complete proteome.
SQ SEQUENCE 240 AA; 27745 MW; 22BD5DEB190BBD CRC64;

Query Match 28.5%; Score 44.5; DB 1; Length 240;
Best Local Similarity 42.9%; Pred. No. 6.7;
Matches 9; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 1 SRXHSMEKXRTPDIN-PAWY 20
DB 197 ARTRLVLMESYTPDFNSPTWY 217

RESULT 8
ID Y587_PASMU STANDARD; PRT; 288 AA.
AC 09CNS6:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein PM0587.
DE PM0587.
GN Pasteurella multocida.
OS Pasteurella multocida.
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;
OC Pasteurella.
OX NCBI_TaxID=747;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=PM70;
RX MEDLINE=21145866; PubMed=11248100;
RA May B.U., Zhang Q., Li L.L., Paustian M.L., Whitlam T.S., Kapur V.;
RA "Complete genomic sequence of Pasteurella multocida PM70."
RA Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).
CC -1- SIMILARITY: BELONGS TO THE FRUCTOSAMINE KINASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
-----

```

DR EMBL: AE006094; AAK02671.1; -  
 KM Hypothetical protein; Transferase; Kinase; Complete proteome.  
 SQ SEQUENCE 288 AA; 33778 MM; FAD2F6C26014D940 CRC64;

Query Match 28.2%; Score 44; DB 1; Length 288;  
 Best Local Similarity 34.8%; Pred. No. 9.7;  
 Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 6 HSMKRTDPINAWYXXGIRPV 28  
 Db 21 HKEKHGEMHEAWIIDGIQPV 43

RESULT 9  
 NER3\_HUMAN STANDARD; PRT; 428 AA.  
 AC Q90049; Q9NOE1;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Stalldase 3 (EC 3.2.1.18) (Membrane stalldase) (Ganglioside stalldase)  
 DE (N-acetyl-alpha-neuraminidase 3).  
 GN NEU3.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=9933533; PubMed=10405317;  
 RA Mada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;  
 RT "Cloning, expression, and chromosomal mapping of a human ganglioside  
 stalldase".  
 RT Biochem. Biophys. Res. Commun. 261:21-27(1999).  
 RL [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Skeletal muscle;  
 RX PubMed=10861246;  
 RA Monti E., Bassi N.T., Papini N., Riboni M., Manzoni M., Venerando B.,  
 Ricci G., Preti A., Ballabio A., Rettamanli G., Borsani G.;  
 RT "Identification and expression of NEU3, a novel human stalldase  
 associated to the plasma membrane".  
 RL Biochem. J. 349:343-351(2000).  
 CC -1- FUNCTION: Plays a role in modulating the ganglioside content of  
 the lipid bilayer at the level of membrane-bound staly  
 glycoconjugates.  
 CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,  
 alpha-(2->8)-glycosidic linkages of terminal sialic residues in  
 oligosaccharides, glycoproteins, glycolipids, colominic acid and  
 synthetic substrates.  
 CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
 CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis,  
 adrenal gland and thymus, followed by pancreas, liver, heart and  
 thymus. Weakly expressed in kidney, placenta, brain and lung.  
 CC -1- MISCELLANEOUS: Optimum pH is 3.8.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
 CC -1- SIMILARITY: CONTAINS 3 BNR REPEATS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AB008185; BAA82611.1; -  
 DR EMBL: Y18563; CAB96131.1; ALT\_INIT.  
 DR MIM: 604617; -  
 DR InterPro: IPR002860; BNR.  
 DR Pfam: PF02012; BNR; 3.

KW Hydrolase; Glycosidase; Membrane; Repeat.  
 FT REPEAT 129 140 BNR 1.  
 FT REPEAT 203 214 BNR 2.  
 FT REPEAT 254 265 BNR 3.  
 FT SITE 24 27 FRIP MOTIF.  
 FT ACT\_SITE 25 25 BY SIMILARITY.  
 FT ACT\_SITE 45 45 POTENTIAL.  
 FT ACT\_SITE 50 50 POTENTIAL.  
 FT ACT\_SITE 87 87 POTENTIAL.  
 FT ACT\_SITE 225 225 POTENTIAL.  
 FT ACT\_SITE 245 245 POTENTIAL.  
 FT ACT\_SITE 340 340 BY SIMILARITY.  
 FT ACT\_SITE 370 370 POTENTIAL.  
 FT ACT\_SITE 387 387 POTENTIAL.  
 SQ SEQUENCE 428 AA; 48252 MM; 35D1DD9359A78C98 CRC64;

Query Match 28.2%; Score 44; DB 1; Length 428;  
 Best Local Similarity 34.8%; Pred. No. 15;  
 Matches 8; Conservative 5; Mismatches 10; Indels 0; Gaps 0;

QY 6 HSMKRTDPINAWYXXGIRPV 28  
 Db 199 HSLMISDLDGVTWHGRIRPM 221

RESULT 10  
 UVRA\_METH STANDARD; PRT; 962 AA.  
 ID UVRA\_METH  
 AC O26543;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Excinuclease ABC subunit A.  
 GN UVRA OR MTH43.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteriales; Methanobacteriaceae;  
 OC Methanothermobacter.  
 OX NCBI\_TaxID=145262;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=DELTA H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Delonghery C., Lee H.-M., Dubois J.,  
 Alledge T., Bashirzadeh R., Blakely D., Cook R., Gilbert R.,  
 Harrison D., Hoang L., Keagle P., Lumm W., Pochler B., Qiu D.,  
 Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 delta: functional analysis and comparative genomics".  
 RL J. Bacteriol. 179:7135-7155(1997).  
 CC -1- FUNCTION: THE ABC EXCISION NUCLEASE IS A DNA REPAIR ENZYME THAT  
 CATALYZES THE EXCISION REACTION OF UV-DAMAGED NUCLEOTIDE SEGMENTS  
 PRODUCING OLIGOMERS HAVING THE MODIFIED BASE(S). UVRA IS AN APASE  
 AND A DNA-BINDING PROTEIN THAT PREFERENTIALLY BINDS SINGLE-  
 STRANDED OR UV-IRRADIATED DOUBLE-STRANDED DNA (BY SIMILARITY).  
 CC -1- SUBUNIT: CONSISTS OF THREE SUBUNTS: UVRA, UVRB AND UVRC.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE ABC TRANSPORTER FAMILY. UVRA SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AE000828; AAB84949.1; -  
 DR EMBL: Y18563; CAB96131.1; ALT\_TRANSPORTR.  
 DR InterPro: IPR003439; ABC\_TRANSPORTR.  
 DR Pfam: PF00005; ABC\_tran; 2.

DR PROSITE: PS00211; ABC\_TRANSPORTER; 2.  
 KW SOS response; Excision nuclease; DNA repair; ATP-binding; Repeat;  
 KW DNA-binding; Zinc-finger; Complete proteome.  
 FT NP\_BIND 38 45 ATP (POTENTIAL).  
 FT NP\_BIND 649 656 ATP (POTENTIAL).  
 FT ZN\_FING 748 774 C4-TYPE.  
 SQ SEQUENCE 962 AA; 108395 MW; 2C0EF7FC41CCD060 CRC64;

Query Match 28.2%; Score 44; DB 1; Length 962;  
 Best Local Similarity 36.8%; Pred. NO. 33;  
 Matches 14; Conservative 0; Mismatches 6; Indels 18; Gaps 2;

OY 11 RPDPINPAMYX-----XRGIRPVGRF 31  
 DB 703 RTRPSNPATYTGVTTHIRLEFAOTPEARKRGYRP-GRF 739

RESULT 11  
 OPS1\_LIMPO STANDARD; PRT; 376 AA.  
 AC P35360;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Lateral eye opsin.  
 OS Limulus polyphemus (Atlantic horseshoe crab).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
 OC Limulidae; Limulus.  
 OX NCBI\_TaxID=6850;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Lateral eye;  
 RX MEDLINE=93317641; PubMed=8327495;  
 RA Smith W.C., Price D.A., Greenberg R.M., Batteille B.-A.;  
 RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
 Limulus polyphemus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
 CC MEDIANE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
 CC LINKED TO CIS-RETINAL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: LATERAL EYE.  
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
 CC BE PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 520 NM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC OPSIN SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L03791; AAA28273.1; -;  
 DR EMBL: L03781; AAA02498.1; -;  
 DR PIR: B48197; B48197.  
 DR GCRDB: GCR\_0584; -;  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR001760; Opsin.  
 DR Pfam: PF00001; 7cm\_1; 1.  
 DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
 DR PROSITE: PS00237; G\_PROTEIN\_RECPT\_FL\_1; 1.  
 DR PROSITE: PS00262; G\_PROTEIN\_RECPT\_FL\_2; 1.  
 DR PROSITE: PS00238; OPSIN; 1.  
 KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
 KW Phosphorylation; G-protein coupled receptor.  
 FT DOMAIN 1 46 EXTRACELLULAR.  
 FT TRANSMEM 47 71 1 (POTENTIAL).  
 FT DOMAIN 72 83 CYTOPLASMIC.

FT TRANSMEM 84 108 2 (POTENTIAL).  
 FT DOMAIN 109 123 EXTRACELLULAR.  
 FT TRANSMEM 124 143 3 (POTENTIAL).  
 FT DOMAIN 144 162 CYTOPLASMIC.  
 FT TRANSMEM 163 186 4 (POTENTIAL).  
 FT DOMAIN 187 210 EXTRACELLULAR.  
 FT TRANSMEM 211 238 5 (POTENTIAL).  
 FT TRANSMEM 239 274 CYTOPLASMIC.  
 FT TRANSMEM 275 298 6 (POTENTIAL).  
 FT DOMAIN 299 306 EXTRACELLULAR.  
 FT TRANSMEM 307 331 7 (POTENTIAL).  
 FT DOMAIN 332 376 CYTOPLASMIC.  
 FT DISULFID 120 197 RETINAL CHROMOPHORE (BY SIMILARITY).  
 FT BINDING 318 318 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 17 17 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 193 193  
 SQ SEQUENCE 376 AA; 42139 MW; CCE401766B806F26 CRC64;

Query Match 27.6%; Score 43; DB 1; Length 376;  
 Best Local Similarity 46.2%; Pred. NO. 18;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 13 PDPINPAMYXRGCI 25  
 DB 40 PPMNPLWYSILGV 52

RESULT 12  
 OPS2\_LIMPO STANDARD; PRT; 376 AA.  
 AC P35361;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ocellar opsin.  
 OS Limulus polyphemus (Atlantic horseshoe crab).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
 OC Limulidae; Limulus.  
 OX NCBI\_TaxID=6850;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Median ocellar;  
 RX MEDLINE=93317641; PubMed=8327495;  
 RA Smith W.C., Price D.A., Greenberg R.M., Batteille B.-A.;  
 RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
 Limulus polyphemus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
 CC MEDIANE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
 CC LINKED TO CIS-RETINAL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: OCELLAR CELLS; MEDIAN OCELLI.  
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
 CC BE PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 530 NM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC OPSIN SUBFAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: L03792; AAA28274.1; -;  
 DR EMBL: L03782; AAA02499.1; -;  
 DR PIR: A48197; A48197.  
 DR GCRDB: GCR\_0585; -;  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR001760; Opsin.

DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCRHHODOPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEP\_FL\_1; 1.  
 DR PROSITE; PS50262; G-PROTEIN\_RECEP\_FL\_2; 1.  
 DR PROSITE; PS00238; OPSIN; 1.  
 KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
 KW Phosphorylation; G-protein coupled receptor.  
 FT DOMAIN 1 46  
 FT TRANSMEM 47 71 1 (POTENTIAL).  
 FT DOMAIN 72 83 1 (POTENTIAL).  
 FT TRANSMEM 84 108 2 (POTENTIAL).  
 FT DOMAIN 109 123 3 (POTENTIAL).  
 FT TRANSMEM 124 143 4 (POTENTIAL).  
 FT DOMAIN 144 162 4 (POTENTIAL).  
 FT TRANSMEM 163 186 5 (POTENTIAL).  
 FT DOMAIN 187 210 5 (POTENTIAL).  
 FT TRANSMEM 211 238 5 (POTENTIAL).  
 FT DOMAIN 239 274 6 (POTENTIAL).  
 FT TRANSMEM 275 298 6 (POTENTIAL).  
 FT DOMAIN 299 306 7 (POTENTIAL).  
 FT TRANSMEM 307 331 7 (POTENTIAL).  
 FT DOMAIN 332 376 7 (POTENTIAL).  
 FT DISULFID 120 197 197  
 FT BINDING 318 318 197  
 FT CARBOHYD 17 17 17  
 FT CARBOHYD 193 193 193  
 SQ SEQUENCE 376 AA; 42111 MW; FA9647C40531CBFB CRC64;

Query Match 27.6%; Score 43; DB 1; Length 376;  
 Best Local Similarity 46.2%; Pred. No. 18;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 13 PDINPAMYXXRGI 25  
 DB 40 PPMNPLWTSILGV 52

RESULT 13  
 CYCR\_CHRYI STANDARD; PRT; 383 AA.  
 AC 082947;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Photosynthetic reaction center cytochrome C subunit precursor.  
 GN PUPC.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Allochromatium.  
 OX NCBI\_TaxID=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=D;  
 RA Corson G.E., Nagashima K.V., Matsuura K., Sakurai Y., Ruwanti W.,  
 RA Qiu H., Allen R., Knäff D.B.;  
 RT "Primary structure of genes encoding light-harvesting and reaction  
 RT center proteins from Chromatium vinosum";  
 RL Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A  
 CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO  
 CC OXIDIZED PRIMARY ELECTRON DONOR.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID  
 CC ANCHOR (BY SIMILARITY).  
 CC -1- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.  
 CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER  
 CC CYTOCHROME C SUBUNITS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB011811; BAA32742.1; -  
 DR HSSP; P07173; 6PRC.  
 DR InterPro; IPR003158; CytC\_RC.  
 DR InterPro; IPR000345; CytC\_heme\_bind.  
 DR Pfam; PF02276; Cytoc\_RC; 1.  
 DR PROSITE; PS00190; CYTOCHROME\_C; 4.  
 KW Electron transport; Photosynthesis; Reaction center; Heme;  
 KW Membrane; Lipoprotein; Duplication; Signal.  
 FT SIGNAL 1 22  
 FT CHAIN 23 383  
 FT LIPID 23 23  
 FT BINDING 107 107  
 FT BINDING 110 110  
 FT METAL 111 111  
 FT BINDING 152 152  
 FT BINDING 155 155  
 FT METAL 156 156  
 FT BINDING 247 247  
 FT BINDING 250 250  
 FT METAL 251 251  
 FT BINDING 307 307  
 FT BINDING 310 310  
 FT METAL 311 311  
 SQ SEQUENCE 383 AA; 41522 MW; 96BCD91F81B9AE7E CRC64;

Query Match 27.6%; Score 43; DB 1; Length 383;  
 Best Local Similarity 36.8%; Pred. No. 19;  
 Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 12 TPDIINPAMYXXRGRPVGR 30  
 DB 263 TPORITAMVIRHVARDINQ 281

RESULT 14  
 YAAJ\_ECOLI STANDARD; PRT; 476 AA.  
 AC P30143;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative transporter yaaJ.  
 GN YAAJ OR B0007.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RA MEDLINE=92334977; PubMed=1630901;  
 RA yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA Isono K., Mizobuchi K., Nakata A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 0-2.4 min region";  
 RL Nucleic Acids Res. 20:3305-3308(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.D.,  
 RA Mau B., Shao Y.;

```

RT "The complete genome sequence of Escherichia coli K-12."
RL Science 277:1453-1474(1997).
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane
CC (potential).
CC -1- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY
CC (SAF). STRONG, TO H. INFLUENZAE H10183.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: D10483; -; NOT_ANNOTATED_CDS.
DR EMBL: AE000111; AAC73118.1; -.
DR EcoGene: EG11555; ynaJ.
DR InterPro: IPR002293; AA_rel-permease_1.
DR InterPro: IPR001463; Na_ala_symp.
DR Pfam: PF01235; Na_ala_symp.1.
DR PRINTS: PS00175; NAALASMPORT.
DR PROSITE: PS00873; NA_ALANINE_SYMP.1.
KW Hypothetical protein; Transmembrane; Inner membrane; Transport;
KW Symport; Complete proteome.
FT TRANSMEM 4 24
FT TRANSMEM 81 101 POTENTIAL.
FT TRANSMEM 141 161 POTENTIAL.
FT TRANSMEM 174 194 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 233 253 POTENTIAL.
FT TRANSMEM 300 320 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 391 411 POTENTIAL.
FT TRANSMEM 414 434 POTENTIAL.
SQ SEQUENCE 476 AA; 51662 MW; 2FE8B2E12E126E63 CRC64;

Query Match 27.6%; Score 43; DB 1; Length 476;
Best Local Similarity 44.4%; Pred. No. 23;
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 14 DIN-----PAMVXXRGI 25
DB 120 DVNGQFRGSPAMVMAAGL 137

RESULT 15
POL_MLVK STANDARD; PRT; 581 AA.
AC P31793;
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 15-DEC-1998 (Rel. 37, Last annotation update)
DE Pol polyprotein [Contains: Protease (EC 3.4.23.-); Reverse
DE transcriptase (EC 2.7.7.49); Ribonuclease H (EC 3.1.26.4)] (Fragment).
GN POL.
OS Radiation murine leukemia virus (strain Kaplan).
OC Viruses; Retroid viruses; Retroviridae; Gammaretrovirus.
OX NCBI_TaxID=31689;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92333703; PubMed=1629969;
RA Poliquin L., Bergeron D., Fortier J.L., Paquette Y., Bergeron R.,
RA Rassart E.;
RT "Determinants of thymotopism in Kaplan radiation leukemia virus and
RT nucleotide sequence of its envelope region.";
RL J. Virol. 66:5141-5146(1992).
CC -1- PTM: SPECIFIC ENZYMAIC CLEAVAGES IN VIVO YIELD NATURE PROTEINS.
CC -1- MISCELLANEOUS: THIS PROTEIN IS SYNTHESIZED AS A GAG-POL
CC POLYPROTEIN.
CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A2; ALSO KNOWN AS THE
CC RETROPEPSIN FAMILY.

```

```

CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: M93052; AAA46525.1; -.
DR PIR: A42743; A42743.
DR HSRP: P29253; IRL.
DR MEROPS: A02.008; -.
DR InterPro: IPR001969; Asp-protease.
DR InterPro: IPR002156; RNaseH.
DR InterPro: IPR001584; Rve.
DR Pfam: PF00075; RNaseH; 1.
DR Pfam: PF00665; rve; 1.
DR PROSITE: PS00141; ASP_PROTEASE; PARTIAL.
KW Hydrolase; Transferase; RNA-directed DNA polymerase;
KW Aspartyl protease; Endonuclease; Polypeptide.
FT NON_TER 1
SQ SEQUENCE 581 AA; 65157 MW; 8D7A38694C8E036E CRC64;

Query Match 27.6%; Score 43; DB 1; Length 581;
Best Local Similarity 34.8%; Pred. No. 29;
Matches 8; Conservative 2; Mismatches 13; Indels 0; Gaps 0;

OY 6 HSMEXRTPDINPAMVXXRGRPV 28
DB 550 HVKATTPPIRPSWVORSONPL 572

```

Search completed: September 13, 2002, 09:11:48  
Job time: 98 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:08:15 : Search time 12.99 Seconds  
(without alignments)  
62.051 Million cell updates/sec

Title: SEQ74-FUSED-TO-SEQ73  
Perfect score: 156  
Sequence: 1 SRHXHSMEXRTPDINPAWYXXRGIRPGRFX 33

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfillst.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	147	94.2	31 3 US-09-105-678A-8	Sequence 8, Appli
2	147	94.2	31 3 US-09-105-678A-9	Sequence 9, Appli
3	147	94.2	31 3 US-09-105-678A-37	Sequence 37, Appli
4	147	94.2	31 3 US-09-105-678A-43	Sequence 43, Appli
5	147	94.2	31 4 US-09-172-353-4	Sequence 4, Appli
6	147	94.2	31 4 US-08-776-971-47	Sequence 47, Appli
7	147	94.2	31 4 US-08-776-971-61	Sequence 61, Appli
8	147	94.2	31 4 US-09-421-208-8	Sequence 8, Appli
9	147	94.2	31 4 US-09-421-208-9	Sequence 9, Appli
10	147	94.2	31 4 US-09-421-208-37	Sequence 37, Appli
11	147	94.2	31 4 US-09-421-208-43	Sequence 43, Appli
12	147	94.2	32 3 US-09-105-678A-38	Sequence 38, Appli
13	147	94.2	32 3 US-09-105-678A-44	Sequence 44, Appli
14	147	94.2	32 4 US-08-776-971-48	Sequence 48, Appli
15	147	94.2	32 4 US-08-776-971-62	Sequence 62, Appli
16	147	94.2	32 4 US-09-421-208-38	Sequence 38, Appli
17	147	94.2	32 4 US-09-421-208-44	Sequence 44, Appli
18	147	94.2	33 3 US-09-105-678A-39	Sequence 39, Appli
19	147	94.2	33 3 US-09-105-678A-45	Sequence 45, Appli
20	147	94.2	33 4 US-08-776-971-49	Sequence 49, Appli
21	147	94.2	33 4 US-08-776-971-63	Sequence 63, Appli
22	147	94.2	33 4 US-09-421-208-39	Sequence 39, Appli
23	147	94.2	33 4 US-09-421-208-45	Sequence 45, Appli
24	147	94.2	83 4 US-08-776-971-45	Sequence 45, Appli
25	147	94.2	83 4 US-08-776-971-124	Sequence 124, App
26	147	94.2	83 4 US-08-776-971-137	Sequence 137, App
27	147	94.2	87 4 US-08-776-971-59	Sequence 59, Appli

28	147	94.2	87 4 US-08-776-971-135	Sequence 135, App
29	147	94.2	87 4 US-08-776-971-138	Sequence 138, App
30	146	93.6	31 3 US-09-105-678A-7	Sequence 7, Appli
31	146	93.6	31 3 US-09-105-678A-31	Sequence 31, Appli
32	146	93.6	31 4 US-08-776-971-5	Sequence 5, Appli
33	146	93.6	31 4 US-08-776-971-97	Sequence 97, Appli
34	146	93.6	31 4 US-09-421-208-7	Sequence 7, Appli
35	146	93.6	31 4 US-09-421-208-31	Sequence 31, Appli
36	146	93.6	32 3 US-09-105-678A-32	Sequence 32, Appli
37	146	93.6	32 4 US-08-776-971-6	Sequence 6, Appli
38	146	93.6	32 4 US-09-421-208-32	Sequence 32, Appli
39	146	93.6	33 3 US-09-105-678A-33	Sequence 33, Appli
40	146	93.6	33 4 US-08-776-971-7	Sequence 7, Appli
41	146	93.6	33 4 US-09-421-208-33	Sequence 33, Appli
42	146	93.6	98 4 US-08-776-971-1	Sequence 1, Appli
43	146	93.6	98 4 US-08-776-971-14	Sequence 44, Appli
44	146	93.6	98 4 US-08-776-971-122	Sequence 122, App
45	146	93.6	98 4 US-08-776-971-131	Sequence 131, App

## ALIGNMENTS

## RESULT 1

US-09-105-678A-8  
Sequence 8, Application US/09105678A

Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESSES:

ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/105,678A

FILING DATE: 26-JUN-1998

PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-105-678A-8

Query Match 94.2% Score 147, DB 3, Length 31;  
Best Local Similarity 83.9% Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
DB 1 SRAHQHSMETRPDINPAMYTGGRIRPVGRF 31

RESULT 2  
US-09-105-678A-9  
; Sequence 9, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-9

Query Match 94.2%; Score 147; DB 3; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
DB 1 SRAHQHSMETRPDINPAMYTGGRIRPVGRF 31

RESULT 3  
US-09-105-678A-37  
; Sequence 37, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-37

Query Match 94.2%; Score 147; DB 3; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
DB 1 SRAHQHSMETRPDINPAMYTGGRIRPVGRF 31

RESULT 4  
US-09-105-678A-43  
; Sequence 43, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.



RESULT 6  
US-08-776-971-47  
Sequence 47, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kikada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 110 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

RESULT 7  
 US-08-776-971-61  
 Sequence 61, Application US/0876971B  
 Patent No. 6228384  
 GENERAL INFORMATION:  
 APPLICANT: Hinuma, Shuji  
 Habata, Yugo  
 Kawamata, Yuji  
 Hosoya, Masaki  
 Fujii, Ryo  
 Fukusumi, Shoji  
 Kitada, Chieko  
 TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
 NUMBER OF SEQUENCES: 140  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: DICE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 STREET: 130 Water Street  
 CITY: Boston  
 STATE: MA  
 COUNTRY: USA  
 ZIP: 02109  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Diskette  
 COMPUTER: IBM compatible

```

:
: OPERATING SYSTEM: DOS
: SOFTWARE: FastSeq for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/776,971B
: FILING DATE: 06-Feb-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP96/03821
: FILING DATE: 28-DEC-1996
: APPLICATION NUMBER: JP 7/34371
: FILING DATE: 28-DEC-1995
: APPLICATION NUMBER: JP 8/59419
: FILING DATE: 15-MAR-1996
: APPLICATION NUMBER: JP 8/21805
: FILING DATE: 12-AUG-1996
: APPLICATION NUMBER: JP 8/246573
: FILING DATE: 18-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 47176
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 61:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: Internal
: SEQUENCE DESCRIPTION: SEQ ID NO: 61:
US-08-776-971-61

Query Match          94.2%; Score 147; DB 4; Length 31;
Best Local Similarity 83.9%; Pred. No. 1.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31
   ||||| ||||| ||||| ||||| |||||
Db 1 SRTHRSMETRTPDINPAMVAGSIRGVGRF 31

RESULT 8
US-09-421-208-8
: Sequence 8, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
```

```

:
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 8:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-09-421-208-8

Query Match          94.2%; Score 147; DB 4; Length 31;
Best Local Similarity 83.9%; Pred. No. 1.3e-17;
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31
   ||||| ||||| ||||| ||||| |||||
Db 1 SRAHQSMETRTPDINPAMVGTGIRPVGRF 31

RESULT 9
US-09-421-208-9
: Sequence 9, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: floppy disk
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
```

; MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 94.2%; Score 147; DB 4; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
|| | ||||| ||||| ||||| ||||| |||||  
Db 1 SRTHRSMETRTPDINPAMYASRGIRPVGRF 31

RESULT 10  
US-09-421-208-37  
; Sequence 37, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 37:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-37

Query Match 94.2%; Score 147; DB 4; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
|| | ||||| ||||| ||||| ||||| |||||  
Db 1 SRAHSHMETRTPDINPAMYTGRIAPVGRF 31

RESULT 11  
US-09-421-208-43  
; Sequence 43, Application US/09421208

; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-43

Query Match 94.2%; Score 147; DB 4; Length 31;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
|| | ||||| ||||| ||||| ||||| |||||  
Db 1 SRTHRSMETRTPDINPAMYASRGIRPVGRF 31

RESULT 12  
US-09-105-678A-38  
; Sequence 38, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3440  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 38:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-38

Query Match 94.2%; Score 147; DB 3; Length 32;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
Db 1 SRAHQSHMETRTPDINPAMYTGRRGIRPVGRF 31

RESULT 13  
US-09-105-678A-44  
Sequence 44, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3440  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 44:

SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-44

Query Match 94.2%; Score 147; DB 3; Length 32;  
Best Local Similarity 83.9%; Pred. No. 1.3e-17;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 SRXHXSMEXRTPDINPAMYXXRGIRPVGRF 31  
Db 1 SRTHRSMETRTPDINPAMYASRGIRPVGRF 31

RESULT 14  
US-08-776-971-48  
Sequence 48, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
APPLICANT: Habata, Yugo  
APPLICANT: Kawamata, Yuji  
APPLICANT: Hosoya, Masaki  
APPLICANT: Fujii, Ryo  
APPLICANT: Fukusumi, Shoji  
APPLICANT: Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3440  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein

```

;      FRAGMENT TYPE: Internal
;      SEQUENCE DESCRIPTION: SEQ ID NO: 48:
US-08-776-971-48

```

Query Match	94.2%	Score 147	DB 4	Length 32
Best Local Similarity	83.9%	Pred. No. 1.3e-17		
Matches	26	Conservative	0	Mismatches 5; Indels 0; Gaps 0;
Qy	1	SRXXHSMEXRTPDINPAMWYXXRCGRPVGR	31	
Db	1	SRAHQHSMTETRPDINPAMWYTGRCGRPVGR	31	

RESULT 15  
US-08-776-971-62  
Sequence 53 Available US/087760715

APPLICANT: Hinuma, Shuji

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

```

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEO for Windows Version 2.0
CURRENT APPLICATION DATA:

```

APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

INFORMATION FOR SEQ ID NO: 62:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids

```

; MOLECULE TYPE: protein
; FRAGMENT TYPE: internal
; SEQUENCE DESCRIPTION: SEQ ID NO: 62:
US-08-776-971-62

```

Query Match	94.2%	Score 147	DB 4	Length 32
Best Local Similarity	83.9%	Pred. No. 1	3e-17	
Matches	26	Conservative	0	Mismatches 5; Indels 0; Gaps 0;
QY	1	SRKHHHSMEXRPPDINPAMYXXRGIRPGRF	31	
DB	1	STHRHSHWEIRTPDINPAMYASRGIRPGRF	31	

Search completed: September 13, 2002, 09:10:33  
Job time: 138 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OK protein - protein search, using sw model

Run on: September 13, 2002, 09:09:05 ; Search time 18.36 Seconds  
(without alignments)  
172.710 Million cell updates/sec

Title: SEQ74-FUSED-TO-SEQ73  
Perfect score: 156  
Sequence: 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRFX 33

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	147	94.2	83	2 JC7607	prolactin-releasin
2	49.5	31.7	664	2 F83376	conserved hypothet
3	49	31.4	220	2 C83292	probable glutathio
4	49	31.4	790	2 T47959	hypothetical prote
5	47	30.1	128	2 S76955	hypothetical prote
6	46	29.5	118	2 AC3169	hypothetical prote
7	46	29.5	314	2 B70569	hypothetical prote
8	46	29.5	333	2 H82852	hydroxybenzoate oc
9	46	29.5	348	2 T21648	hypothetical prote
10	46	29.5	798	2 S11210	probable unr prote
11	46	29.5	1236	2 T50904	Mg protoporphylin
12	46	29.5	1292	2 T31462	probable magnesium
13	45	28.8	240	2 B75318	ferrityochelin-bin
14	45	28.8	501	2 T48336	hypothetical prote
15	45	28.8	637	2 D87559	sensory box histid
16	45	28.8	767	2 T21699	hypothetical prote
17	44.5	28.5	240	2 D64688	probable 1-acylgly
18	44.5	28.5	954	2 A83652	hypothetical prote
19	44	28.2	72	2 E91002	probable regulator
20	44	28.2	159	2 G82669	ubiquinone biosynt
21	44	28.2	250	2 G83400	hypothetical prote
22	44	28.2	284	2 F71015	hypothetical prote
23	44	28.2	548	2 T47548	hypothetical prote
24	44	28.2	962	2 H69157	exonuclease ABC c
25	43.5	27.9	503	2 A82193	Sun/nucleolar prot
26	43.5	27.9	852	2 T28790	hypothetical prote
27	43.5	27.9	1501	2 T45623	hypothetical prote
28	43	27.6	232	2 G75608	hypothetical prote
29	43	27.6	309	2 T32376	hypothetical prote

30	43	27.6	342	2 F91007	probable gentisate
31	43	27.6	342	2 H85851	probable 1,2-dioxy
32	43	27.6	376	2 A48197	opsin, ocular - A
33	43	27.6	376	2 B48197	opsin, lateral eye
34	43	27.6	455	2 D70885	probable aldC prot
35	43	27.6	476	2 G64720	probable amino aci
36	43	27.6	476	2 G90629	probable inner mem
37	43	27.6	476	2 G85480	inner membrane tra
38	43	27.6	581	2 A42743	pol polypotelein -
39	43	27.6	719	2 S61046	ARPI protein - yea
40	43	27.6	843	1 GNVWK	pol polypotelein -
41	43	27.6	1196	1 GNMVGV	pol polypotelein -
42	43	27.6	1196	1 GNMVGV	pol polypotelein -
43	43	27.6	1296	2 T16859	hypothetical prote
44	42.5	27.2	303	2 AH2016	hypothetical prote
45	42.5	27.2	443	2 T21499	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces archidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PIRP  
A:Introns: 33/1

Query Match 94.2%; Score 147; DB 2; Length 83;  
Best Local Similarity 83.9%; Pred. No. 2.5e-16;  
Matches 26; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 SRXHXSMEXRTPDINPAWYXXRGIRPVGRF 31  
DB 22 SRAHQSMETRPDINPAWYTGIRPVGRF 52

RESULT 2  
F83376  
conserved hypothetical protein PA2151 [Imported] - Pseudomonas aeruginosa (strain PAO  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83376  
R:Stover, C.K.; Pham, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adam, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lapid, K.; L.:; Lory, S.; Olson, M.V.  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa  
A:Reference number: AB2950; MUID:20437337  
A:Accession: F83376  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STO>  
A:Cross-references: GB:AE004642; GB:AE004091; NID:99948163; PIDN:AG05539.1; GSPDB:GN  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA2151

Query Match 31.7%; Score 49.5; DB 2; Length 664;  
Best Local Similarity 62.5%; Pred. No. 8.1;  
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 12 TPDINPAMYXXRGIRP 27  
| | | | | : | | |  
| | | | | : | | |  
DB 478 TPDINP-WFQIRSGRP 492

RESULT 3

Probable glutathione S-transferase PA2821 [Imported] - Pseudomonas aeruginosa (strain PA  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001  
C:Accession: C63292  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Bt  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: AB2950; MUID:20437337  
A:Accession: C63292  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-220 <STO>  
A:Cross-references: GB:AE004709; GB:AE004091; NID:99948904; PIDN:AAG06209.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2821  
C:Superfamily: plaiace glutathione transferase

Query Match 31.4%; Score 49; DB 2; Length 220;  
Best Local Similarity 44.0%; Pred. No. 3.1;  
Matches 11; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

OY 6 HSMEXRTPDINPAMYXXRGIRPVG 30  
: : | | | | | : | | |  
DB 28 YOLEIAVPGOPAWY--REISPLGR 50

RESULT 4

hypothetical protein F15G16.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quet  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z24480  
A:Accession: T47959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEH>  
A:Cross-references: EMBL:AL132959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 39/1; 678/2; 698/3; 773/2  
A:Note: F15G16.60

Query Match 31.4%; Score 49; DB 2; Length 790;  
Best Local Similarity 47.4%; Pred. No. 12;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 13 PDINPAMYXXRGIRPVG 31  
| | | | | : | | |  
| | | | | : | | |  
DB 366 PPHNPTVGSRLQPHGRW 384

RESULT 5

S76955  
hypothetical protein - Synechocystis sp. (strain PCC 6803)  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76955  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
S.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76955  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-128 <KAN>  
A:Cross-references: EMBL:D90917; GB:AB001339; NID:g1653836; PIDN:BA018867.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 30.1%; Score 47; DB 2; Length 128;  
Best Local Similarity 33.3%; Pred. No. 3.8;  
Matches 11; Conservative 5; Mismatches 13; Indels 4; Gaps 1;

OY 1 SRXHXSMEXRTPDINPAMYXXR---GIRPVG 29  
| : | : | : | : | : | | | | |  
DB 11 SPENNOSIDCSSPDPSSRKWYQARLFSGITPYG 43

RESULT 6

hypothetical protein Atu5077 [Imported] - Agrobacterium tumefaciens (strain C58, Dupo  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AC3169  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Moo  
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Bliddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AC3169  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-118 <KUR>  
A:Cross-references: GB:AE008687; PIDN:RAL45769.1; PID:g17743503; GSPDB:GN00188  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu5077  
A:Genome: plasmid

Query Match 29.5%; Score 46; DB 2; Length 118;  
Best Local Similarity 34.4%; Pred. No. 5;  
Matches 11; Conservative 4; Mismatches 13; Indels 4; Gaps 1;

OY 2 RXHXSMEXRTPDI---NPAMYXXRGIRPVG 29  
: : | : | : | : | : | : |  
DB 19 KLHKHNAARRRDPDIGARVQPATAPHRGLAGIG 50

RESULT 7

hypothetical protein RV3485c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70569  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998



A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987  
A:Accession: B70569  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-314 <CDL>  
A:Cross-references: GB:295390; GB:AL123456; NID:93261766; PIDN:CA808708.1; PID:g2104408  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV3485c  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
F:46-227/Domain: short-chain alcohol dehydrogenase homology <SADH>

Query Match 29.5%; Score 46; DB 2; Length 314;  
Best Local Similarity 31.8%; Pred. No. 14;  
Matches 7; Conservative 4; Mismatches 11; Indels 0; Gaps 0;

OY 6 HSMEXRTPDINPAMYXRGIRP 27  
| : : : : :  
Db 205 HMKRLADELGPWSVRVNSIRP 226

RESULT 8  
H82852  
hydroxybenzoate octaprenyltransferase Xf0068 [Imported] - Xylella fastidiosa (strain 9a5  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: H82852  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82852  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <STM>  
A:Cross-references: GB:AE003860; GB:AE003849; NID:9104830; PIDN:AAE82881.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Brites, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carver, H  
as-Melo, E.; Docena, C.; El-Dorri, H.; Facinanci, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; From  
J.D.; Jungella, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, B  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsunako, M.H.; Valada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: Xf0068  
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 29.5%; Score 46; DB 2; Length 333;  
Best Local Similarity 53.3%; Pred. No. 15;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 15 INPAMYXXRGIRPVG 29  
: : : : :  
Db 54 LDPYMKLARGDPRVG 68

RESULT 9  
T21648  
hypothetical protein F32B6.3 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans

C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T21648  
R:Basham, V.  
submitted to the EMBL Data Library, October 1996

A:Reference number: Z19453  
A:Accession: T21648  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-348 <MTL>  
A:Cross-references: EMBL:Z81074; PIDN:CA803039.1; GSPDB:GN00022; CESP:F32B6.3  
A:Experimental source: clone F32B6  
C:Genetics:  
A:Gene: CESP:F32B6.3  
A:Map position: 4  
A:Introns: 56/3; 105/1; 146/3; 185/1; 223/1; 258/2; 320/3

Query Match 29.5%; Score 46; DB 2; Length 348;  
Best Local Similarity 38.9%; Pred. No. 15;  
Matches 7; Conservative 2; Mismatches 9; Indels 0; Gaps 0;

OY 2 RXHXSMEKRPDINPAW 19  
| : : : : :  
Db 112 RKRHQLLELAQPDLENGW 129

RESULT 10  
S11210  
probable unr protein - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 21-Nov-1993 #sequence\_revision 10-Nov-1995 #text\_change 05-Nov-1999  
C:Accession: S11210  
R:Jeffers, M.; Pacinuci, R.; Pellicer, A.  
Nucleic Acids Res. 18, 4891-4899, 1990  
A:Title: Characterization of unr; a gene closely linked to N-ras.  
A:Reference number: S11210; MUID:90370473  
A:Accession: S11210  
A:Molecule type: mRNA  
A:Residues: 1-798 <JREF>  
A:Cross-references: EMBL:X52311; NID:957454; PIDN:CAJ36549.1; PID:957455  
C:Keywords: DNA binding

Query Match 29.5%; Score 46; DB 2; Length 798;  
Best Local Similarity 39.1%; Pred. No. 36;  
Matches 9; Conservative 4; Mismatches 10; Indels 0; Gaps 0;

OY 6 HSMEXRTPDINPAMYXRGIRPV 28  
| : : : : :  
Db 583 HSNVGTTEANPTIYSGKVRPL 605

RESULT 11  
T50904  
Mg protoporphyrin methyl transferase [Imported] - Rubrivivax gelatinosus  
C:Species: Rubrivivax gelatinosus  
C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 02-Sep-2000  
C:Accession: T50904  
R:Nagashima, K.V.; Igarashi, N.; Harada, J.; Nagashima, S.; Matsuura, K.; Shmada, K.  
submitted to the EMBL Data Library, November 1999  
A:Description: Determination of Nucleotide Sequences of Rubrivivax gelatinosus Photos  
A:Reference number: 225270  
A:Accession: T50904  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-1236 <NAG>  
A:Cross-references: EMBL:AB034704; PIDN:BA94057.1  
A:Experimental source: strain IL144  
C:Genetics:  
A:Gene: bchh  
C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase

Query Match 29.5%; Score 46; DB 2; Length 1236;  
 Best Local Similarity 32.1%; Pred. No. 56;  
 Matches 9; Conservative 5; Mismatches 8; Indels 6; Gaps 1;  
 OY 7 SMEXRTPDINPAMY-----XXRGIRPV 28  
 DB 1116 ALETRRLNPKWYEGMKHGYEGVQR 1143

RESULT 12  
 T31462  
 probable magnesium chelatase (EC 4.99.1.-) chain H BchH - Hellobacillus mobilis  
 C:Species: Hellobacillus mobilis  
 C:Date: 02-Sep-2000 #sequence\_revision 02-Sep-2000 #text\_change 15-Sep-2000  
 C:Accession: T31462  
 R:Xiong, J.; Inoue, K.; Bauer, C.E.  
 Proc. Natl. Acad. Sci. U.S.A. 95, 14851-14856, 1998  
 A:Title: Tracking molecular evolution of photosynthesis by characterization of a major F  
 A:Reference number: 221036; MUID:99061957  
 A:Accession: T31462  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-1292 <XIO>  
 A:Cross-references: EMBL:AF080002; NID:g3820536; PID:g3820560; PIDN:AAC84033.1  
 C:Genetics:  
 A:Gene: bchH  
 C:Superfamily: Rhodobacter capsulatus magnesium-protoporphyrin O-methyltransferase  
 C:Keywords: lyase

Query Match 29.5%; Score 46; DB 2; Length 1292;  
 Best Local Similarity 34.6%; Pred. No. 59;  
 Matches 9; Conservative 3; Mismatches 8; Indels 6; Gaps 1;  
 OY 9 EXRTPDINPAMY-----XXRGIRPV 28  
 DB 1176 ETRRTKLNPKWYEGMKHGYEGVQR 1201

RESULT 13  
 B75318  
 ferritychoelin-binding protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 17-Mar-2000  
 C:Accession: B75318  
 R:White, O.; Eisen, J.A.; Heidelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.  
 Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896  
 A:Accession: B75318  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-240 <WH>  
 A:Cross-references: GB:AE002044; GB:AE000513; NID:g6459872; PIDN:AAF11635.1; PID:g645988  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DR2089  
 A:Map position: 1

Query Match 28.8%; Score 45; DB 2; Length 240;  
 Best Local Similarity 32.0%; Pred. No. 15;  
 Matches 8; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 6 HSMEXRTPDINPAMYXXRGIRPVGR 30  
 DB 63 HTLEDLTPDIRPTAFVAPSADYIQ 87

RESULT 14  
 T48336

hypothetical protein F15A17.180 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cress)  
 C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T48336  
 R:Bayan, M.; Terry, N.; Ardiles, W.; Buysheert, C.; Dasseville, R.; De Clerck, R.;  
 ewes, H.W.; Rudd, S.; Lemcke, K.; Mayer, K.F.X.  
 submitted to the Protein Sequence Database, April 2000  
 A:Reference number: 224491  
 A:Accession: T48336  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-501 <BE>  
 A:Cross-references: EMBL:AL163002  
 A:Experimental source: cultivar Columbia; BAC clone F15A17  
 C:Genetics:  
 A:Map position: 5  
 A:Introns: 63/1; 194/2; 280/1  
 A:Note: F15A17.180

Query Match 28.8%; Score 45; DB 2; Length 501;  
 Best Local Similarity 42.9%; Pred. No. 32;  
 Matches 6; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 4 HXSMEXRTPDINP 17  
 DB 24 HLHHLQQQIPDLP 37

RESULT 15  
 D87559  
 sensory box histidine kinase/response regulator [imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
 C:Accession: D87559  
 R:Nierman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
 B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
 n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A:Title: Complete genome sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: D87559  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-637 <STO>  
 A:Cross-references: GB:AE005673; NID:g13424056; PIDN:AAK24472.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC2501

Query Match 28.8%; Score 45; DB 2; Length 637;  
 Best Local Similarity 44.0%; Pred. No. 41;  
 Matches 11; Conservative 3; Mismatches 11; Indels 0; Gaps 0;

OY 4 HXSMEXRTPDINPAMYXXRGIRPV 28  
 DB 22 HRDSDRLRSPAINPAIRVRLRAV 46

Search completed: September 13, 2002, 09:10:59  
 Job time: 114 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:11:35 ; Search time 311.85 Seconds

(without alignments)  
17.197 Million cell updates/sec

Title: US-09-446-543a-61

Perfect score: 171  
Sequence: 1 SRTHSHMEIRTPDINPAMVASRGIRPVGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

SPRTEMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organella:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	98	57.3	117	13	Q9W624
2	56.5	33.0	664	16	Q911W4
3	55	32.2	315	2	087474
4	54	31.6	790	10	Q9M371
5	53.5	31.3	314	16	006348
6	53	31.0	327	10	Q94K09
7	53	31.0	327	10	Q94E29
8	53	31.0	637	16	Q9A5E9
9	52	30.4	785	17	026276
10	51.5	30.1	420	17	Q931Z7
11	51	29.8	54	4	Q9UJF9
12	51	29.8	277	11	Q9D8A2
13	51	29.8	286	11	Q9DAE5
14	51	29.8	450	11	Q9DA19
15	51	29.8	455	4	060687
16	50	29.2	123	2	Q45883

17	50	29.2	128	16	P74747	P74747 synechocyst
18	50	29.2	692	2	Q9L8J6	Q9L8J6 rhodospirill
19	50	29.2	939	5	Q9G8A0	Q9G8A0 hemiceutrot
20	49.5	28.9	176	3	Q08689	Q08689 saccharomyc
21	49	28.7	72	2	Q9EYB3	Q9EYB3 escherichia
22	49	28.7	173	17	Q97A00	Q97A00 thermoplasm
23	49	28.7	326	2	Q9F642	Q9F642 stigmatella
24	48.5	28.4	1501	10	Q9SD86	Q9SD86 arabidopsis
25	48	28.1	220	16	Q91022	Q91022 pseudomonas
26	48	28.1	335	2	088002	088002 bordetella
27	48	28.1	335	2	Q45375	Q45375 bordetella
28	48	28.1	348	5	Q45431	Q45431 caenorhabdi
29	48	28.1	455	16	Q33340	Q33340 mycobacteri
30	48	28.1	514	5	Q961A3	Q961A3 drosophila
31	48	28.1	596	5	Q9VP62	Q9VP62 drosophila
32	48	28.1	863	5	Q9RVV3	Q9RVV3 drosophila
33	48	28.1	1084	5	Q9TYW4	Q9TYW4 caenorhabdi
34	47.5	27.8	333	2	Q9RUI0	Q9RUI0 streptomyce
35	47.5	27.8	345	16	Q9K8V2	Q9K8V2 bacillus ha
36	47.5	27.8	503	16	Q9KRY1	Q9KRY1 vibrio chol
37	47.5	27.8	805	5	Q9TGS3	Q9TGS3 drosophila
38	47	27.5	145	5	Q95GB5	Q95GB5 leishmania
39	47	27.5	159	16	Q9PD41	Q9PD41 xyella fas
40	47	27.5	215	16	Q9A522	Q9A522 caulobacter
41	47	27.5	284	17	Q50128	Q50128 pyrococcus
42	47	27.5	313	16	Q9PH76	Q9PH76 xyella fas
43	47	27.5	432	10	Q9EYX0	Q9EYX0 arabidopsis
44	47	27.5	501	10	Q9LYX0	Q9LYX0 arabidopsis
45	47	27.5	881	5	Q9NE42	Q9NE42 leishmania

## ALIGNMENTS

RESULT	ID	Q9W624	PRELIMINARY;	PRT;	117 AA.
AC	Q9W624	01-NOV-1999 (TREMUREL. 12, Created)			
DT	01-NOV-1999	(TREMUREL. 12, Last sequence update)			
DT	01-DEC-2001	(TREMUREL. 19, Last annotation update)			
DE	C-RF AMIDE.				
OS	Carassius auratus (Goldfish).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;				
OC	Cypriniformes; Cyprinidae; Carassius.				
OX	NCBI_TaxID=7957;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE-BRAIN;				
RA	Satoh H., Minakata H., Fujimoto M.;				
RT	"Carassius Rhamdi (C-RF amide)."				
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.				
DR	EMBL: AB020024; BAA76662.1; -				
SQ	SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;				

Query Match	57.3%	Score 98;	DB 13;	Length 117;
Best Local Similarity	53.8%	Pred. No. 4.4e-07;		
Matches 14;	Conservative	8;	Mismatches 4;	Indels 0;
Gaps	0;			
QY	6	HSMEIRTPDINPAMVASRGIRPVGRF 31		
DB	50	HVNDNRSEIDPFYVGRVPIGRF 75		
RESULT	2			
Q911W4		PRELIMINARY;	PRT;	664 AA.
AC	Q911W4	01-MAR-2001 (TREMUREL. 16, Created)		
DT	01-MAR-2001	(TREMUREL. 16, Last sequence update)		
DT	01-OCT-2001	(TREMUREL. 18, Last annotation update)		

```

DE  HYPOTHETICAL PROTEIN PA2151.
GN  PA2151.
OS  Pseudomonas aeruginosa.
OC  Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC  Pseudomonas.
OX  NCBI_TaxID=287;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=ATCC 15692 / PA01;
RX  MEDLINE=20437337; PubMed=10984043;
RA  Stoyer C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
RA  Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA  Garner R.L., Goltz L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
RA  Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,
RA  Smith K.A., Spencer D.H., Wong G.R.-S., Wu Z., Paulsen I.T.,
RA  Reiker J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT  Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT  opportunistic pathogen."
RL  Nature 406:959-964(2000).
DR  EMBL; AE004642; AAC05539.1; -.
DR  InterPro; IPR001589; Actinln_act_bind.
DR  InterPro; IPR000461; Alpha_amyase.
DR  Pfam; PF00128; alpha_amyase; 1.
DR  PROSITE; PS00019; ACTININ_1; UNKNOWN_1.
KW  Hypothetical protein; Complete proteome.
SQ  SEQUENCE 664 AA; 76329 MW; 8F59FED54C308AD CRC64;

Query Match 33.0%; Score 56.5; DB 16; Length 664;
Best Local Similarity 45.8%; Pred. No. 5;
Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

QY  4 HRHSMETRPDINPAWYASRGIRP 27
    : : : : : : : : : : : :
DB  470 YRPNEFVNTPDINP-WLQSRGP 492

RESULT 3
ID  087474 PRELIMINARY; PRT; 315 AA.
AC  087474;
DT  01-NOV-1998 (TREMBLrel. 08, Created)
DT  01-NOV-1998 (TREMBLrel. 08, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  TRIMETHOXYTOLUENE OXYGENASE.
GN  DNTD.
OS  Burkholderia cepacia (Pseudomonas cepacia).
OC  Bacteria; Proteobacteria; beta subdivision; Burkholderia group;
OC  Burkholderia.
OX  NCBI_TaxID=292;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=DNT;
RX  MEDLINE=93194809; PubMed=8449889;
RA  Suen W.C., Spain J.C.;
RT  Cloning and characterization of Pseudomonas sp. strain DNT genes for
RT  2,4-dinitrotoluene degradation."
RL  J. Bacteriol. 175:1831-1837(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN=DNT;
RX  MEDLINE=99121037;
RA  Haigler B.E., Johnson G.R., Suen W.C., Spain J.C.;
RT  Biochemical and genetic evidence for meta-ring cleavage of 2,4,5-
RT  trinitroxytoluene in Burkholderia sp. strain DNT."
RL  J. Bacteriol. 181:965-972(1999).
DR  EMBL; AF076848; AAD12738.1; -.
DR  InterPro; IPR000486; Extradiol_dioxygenase.
DR  InterPro; IPR004360; Gly_bio_diox.
DR  Pfam; PF00903; Glyoxalase; 2.
DR  ProDom; PD000977; Extradiol_dioxygenase; 1.
SQ  SEQUENCE 315 AA; 34745 MW; E99261179022961E CRC64;

```

```

Query Match 32.2%; Score 55; DB 2; Length 315;
Best Local Similarity 44.4%; Pred. No. 3; 7;
Matches 16; Conservative 3; Mismatches 11; Indels 6; Gaps 2;

QY  1 SRRHSMETRPD---INPAWYASRGIRP---VGR 30
    : : : : : : : : : : : :
DB  216 SRVHCSFEIODLDAQFLGNKWLASGRKPGWGVGR 251

RESULT 4
ID  09M371 PRELIMINARY; PRT; 790 AA.
AC  09M371;
DT  01-OCT-2000 (TREMBLrel. 15, Created)
DT  01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  HYPOTHETICAL 87.4 KDA PROTEIN.
GN  F15616.60.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA  Mayer K.F.X., Quetler F., Salanoubat M.;
RL  Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  EU Arabidopsis sequencing project;
RL  Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL; AL132959; CAB71097.1; -.
KW  Hypothetical protein.
SQ  SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 31.6%; Score 54; DB 10; Length 790;
Best Local Similarity 52.6%; Pred. No. 14;
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY  13 PDINPAWYASRGIRPVGRF 31
    : : : : : : : : : : : :
DB  366 PPHNPTYSRGLQPHGRW 384

RESULT 5
ID  006348 PRELIMINARY; PRT; 314 AA.
AC  006348;
DT  01-JUL-1997 (TREMBLrel. 04, Created)
DT  01-JUL-1997 (TREMBLrel. 04, Last sequence update)
DT  01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE  HYPOTHETICAL 33.2 KDA PROTEIN.
GN  RV3485C OR MTCY13E12.38C.
OS  Mycobacterium tuberculosis.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX  NCBI_TaxID=1773;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN=H37RV;
RX  MEDLINE=98295987; PubMed=9634230;
RA  Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA  Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,
RA  Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA  Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,
RA  Horsby T., Jagsels K., Krogh A., McLean J., Moule S., Murphy L.,
RA  Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA  Rutter S., Seeger K., Skelton S., Squares S., Squares R.,
RA  Sulton J.E., Taylor K., Whitehead S., Barrett B.G.;
RT  Deciphering the biology of Mycobacterium tuberculosis from the

```

RT complete genome sequence."  
RL Nature 393:537-544(1998).  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
DR EMBL: Z95390; CAB08708.1; -.  
DR HSSP: P29132; 1DFT.  
DR Tuberculin: RV3485C; -.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF01016; adh\_short.1.  
DR PROSITE: PS00061; ADH\_SHORT; UNKNOWN.1.  
KW Complete proteome; Hypothetical protein; Oxidoreductase.  
SQ SEQUENCE 314 AA; 33194 MW; 26144BA917E09274 CRC64;

Query Match 31.3%; Score 53.5; DB 16; Length 314;  
Best Local Similarity 27.5%; Pred. No. 6.2;  
Matches 11; Conservative 5; Mismatches 11; Indels 13; Gaps 1;

OY 1 SRTHR-----HSMEIRTPDINPAWASRGIRP 27  
Db 187 SNTHRMGAYGVTKSAVDHMKLADELGPSSWVRVNSIRP 226

RESULT 6  
O94KU9 PRELIMINARY; PRT; 327 AA.  
AC O94KU9;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE PLASTID-LIPID ASSOCIATED PROTEIN PAP1.  
OS Brassica campestris (field mustard).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21249173; PubMed=11351096;  
RA Kim H.U., Wu S.S., Ratnayake C., Huang A.H.;  
RT "Brassica rapa Has Three Genes That Encode Proteins Associated with  
RT Plant Physiol. 126:330-341(2001).  
RL EMBL: AF290563; AAK57561.1; -.  
DR SEQUENCE 327 AA; 35678 MW; 8C7B87FCD6C02422 CRC64;

Query Match 31.0%; Score 53; DB 10; Length 327;  
Best Local Similarity 43.8%; Pred. No. 7.7;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 4 HRHSMETRPDINPAW 19  
Db 45 HRHDFKVRASDVNDW 60

RESULT 7  
O94FZ9 PRELIMINARY; PRT; 327 AA.  
AC O94FZ9;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE PLASTID-LIPID ASSOCIATED PROTEIN PAP1.  
OS Brassica campestris (field mustard).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Brassica.  
OX NCBI\_TaxID=3711;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21249173; PubMed=11351096;  
RA Kim H.U., Wu S.S., Ratnayake C., Huang A.H.;

RT "Brassica rapa Has Three Genes That Encode Proteins Associated with  
RT Different Neutral Lipids in Plastids of Specific Tissues."  
RL Plant Physiol. 126:330-341(2001).  
DR EMBL: AF290566; AAK57564.1; -.  
SQ SEQUENCE 327 AA; 35644 MW; 6116E7F1B6C02C88 CRC64;

Query Match 31.0%; Score 53; DB 10; Length 327;  
Best Local Similarity 43.8%; Pred. No. 7.7;  
Matches 7; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 4 HRHSMETRPDINPAW 19  
Db 45 HRHDFKVRASDVNDW 60

RESULT 8  
O9A5E9 PRELIMINARY; PRT; 637 AA.  
AC O9A5E9;  
DT 01-JUN-2001 (TReMBLrel. 17, Created)  
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE SENSOR BOX HISTIDINE KINASE/RESPONSE REGULATOR.  
GN CC2501.  
OS Caulobacter crescentus-  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Niernan W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
RA DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
RA Uterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
CC -1- SIMILARITY: THE N-TERMINAL REGION IS SIMILAR TO THAT OF OTHER  
CC REGULATORY COMPONENTS OF SENSOR TRANSDUCTION SYSTEMS.  
CC -1- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.  
DR EMBL: AE005918; AAK24472.1; -.  
DR HSSP: O56312; 3TMV.  
DR TIGR: CC2501.  
DR InterPro: IPR004358; BCTRLSENSOR.  
DR InterPro: IPR003594; HATPase\_C.  
DR InterPro: IPR003661; His\_kin.  
DR InterPro: IPR004359; HIS\_KIN\_slg.  
DR InterPro: IPR001610; PAC.  
DR InterPro: IPR000014; PAS.  
DR InterPro: IPR000700; PAS-assoc\_C.  
DR InterPro: IPR001789; Response\_reg.  
DR Pfam: PF02518; HATPase\_C.1.  
DR Pfam: PF00785; PAC.1.  
DR Pfam: PF00072; response\_reg.1.  
DR Pfam: PF00312; signal.1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR SMART: SM00387; HATPase\_C.1.  
DR SMART: SM00388; Hlska.1.  
DR SMART: SM00086; PAC.1.  
DR SMART: SM00091; PAS.1.  
DR SMART: SM00448; REC.1.  
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;  
KW Transferrase.  
SQ SEQUENCE 637 AA; 68511 MW; 0EDEBAF76FPA8611 CRC64;

Query Match 31.0%; Score 53; DB 16; Length 637;  
Best Local Similarity 48.0%; Pred. No. 16;

Matches	12;	Conservative	4;	Mismatches	9;	Indels	0;	Gaps	0
OY	4	HRHSEIRTPDINPAWYASRGIRPV	28						
DB	22	HRSDDLKSPAINPAIRVRIILRAV	46						
RESULT	9								
ID	026276	PRELIMINARY:	PRT:	785 AA.					
AC	026276.								
DT	01-JAN-1998	(TREMBLrel. 05, Created)							
DT	01-JAN-1998	(TREMBLrel. 05, Last sequence update)							
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)							
DE	SENSOR TRANSDUCTION HISTIDINE KINASE.								
GN	MTN174.								
OS	Methanothermobacter thermotrophicus.								
OC	Archaea, Euryarchaeota; Methanobacteriales; Methanobacteriaceae;								
OC	Methanothermobacter.								
OX	NCBI_TaxID=145262;								
RN	(1)								
RP	SEQUENCE FROM N.A.								
RC	STRAIN-DELTA H:								
RX	MEDLINE=968037514; PubMed=9371463;								
RA	Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,								
RA	Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,								
RA	Harrison D., Hoang L., Keagle P., Lunn W., Potlter B., Qiu D.,								
RA	Safedeva R., Vleare R., Wang Y., Wierzbowski J., Gibson R.,								
RA	Jivani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,								
RA	McConnell S., Shimer G., Goyal A., Pietrovski S., Church G.M.,								
RA	Daniels C.J., Mao J.-I., Rice P., Nollig J., Reeve J.N.;								
RT	*Complete genome sequence of Methanothermobacterium thermoautotrophicum								
RT	deltah: functional analysis and comparative genomics.*;								
RL	J. Bacteriol. 179:7135-7155(1997).								
DR	EMBL; AE000805; AAB84680.1; "								
DR	InterPro: IPR000515; BPD_transp.								
DR	InterPro: IPR003594; HATPase_C.								
DR	InterPro: IPR004359; HTS_KIN_sig.								
DR	InterPro: IPR001610; PAC.								
DR	InterPro: IPR000014; PAS.								
DR	InterPro: IPR000700; PAS-assoc_C.								
DR	Pfam: PF02518; HATPase_C; 1.								
DR	Pfam: PF00785; PAC; 1.								
DR	Pfam: PF00989; PAS; 3.								
DR	Pfam: PF00517; signal; 1.								
DR	SMART; SM00387; HATPase_C; 1.								
DR	SMART; SM00086; PAC; 1.								
DR	SMART; SM00091; PAS; 3.								
DR	PROSITE; PS00402; BPD_TRANSP_INN_MEMBER; UNKNOWN_1.								
SO	Complete proteome.								
SW	SEQUENCE 785 AA; 87726 MW; ADD502C928307986 CRC64;								
Query Match	30.4%;	Score 52;	DB 17;	Length 785;					
Best Local Similarity	46.4%;	Pred. NO. 28;							
Matches	13;	Conservative	1;	Mismatches	12;	Indels	2;	Gaps	1.
OY	4	HRHSEIRTPDINPAW--YASRGIRPV	29						
DB	412	HRATIRRPDGNRYWETVDRPIRTDG	439						
RESULT	10								
ID	093127	PRELIMINARY:	PRT:	420 AA.					
AC	093127.								
DT	01-DEC-2001	(TREMBLrel. 19, Created)							
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)							
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)							
DE	CHAIN LENGTH FACTOR-LIKE PROTEIN.								
GN	AUR2B.								
OS	Streptomyces aureofaciens.								
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;								

OC	Actinomyces:	Streptomyces;	Streptomyces;
OX	NCGI_TaxId=1894;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN-CCM3239;		
RA	Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.,		
RT	"Cloning and characterization of a new polyketide gene cluster in		
RL	Streptomyces aureofaciens CCM3239."		
DR	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.		
SQ	EMBL; AY033994; AAK617719.1; "		
	SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;		
OY	Query Match	30.1%; Score 51.5; DB 2; Length 420;	
	Best Local Similarity	57.9%; Pred. No. 17;	
	Matches 11; Conservative	2; Mismatches 1; Indels 5; Gaps 1;	
Db	18 AWVAS----RGIRPVGRF 31    :      37 AMMAAVLRGSGIRPVGRF 55		
RESULT 11			
O9UJF9	PRELIMINARY;	PRT; 54 AA.	
AC O9UJF9;			
DT 01-MAY-2000 (TREMBLrel. 13, created)			
DT 01-MAY-2000 (TREMBLrel. 13, last sequence update)			
DT 01-MAY-2000 (TREMBLrel. 13, last annotation update)			
DE D147937.3 (SUSHI-REPEAT PROTEIN (SRPUL)) (FRAGMENT).			
GN D147937.3.			
OS Homo sapiens (Human).			
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX NCGI_TaxId=9606;			
RN [1]			
RP SEQUENCE FROM N.A.			
RA Laylor S.;			
RL Submitted (SEP-1999) to the EMBL/Genbank/DBJ databases.			
DR EMBL; AL035608; CAB55682.1; -			
FT NON_TER 54 54			
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;			
OY 12 TPDIHPAWYASRGIRP 27   :      18 TPAVPTWYAGSGIRP 33			
RESULT 12			
O9D8A2	PRELIMINARY;	PRT; 277 AA.	
AC O9D8A2;			
DT 01-JUN-2001 (TREMBLrel. 17, created)			
DT 01-JUN-2001 (TREMBLrel. 17, last sequence update)			
DT 01-JUN-2001 (TREMBLrel. 17, last annotation update)			
DE I700023B02RIK PROTEIN. 1700023B02RIK.			
OS Mus musculus (Mouse).			
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX NCGI_TaxId=10090;			
RN [1]			
RP SEQUENCE FROM N.A.			
RC STRAIN=C57BL/6J; TISSUE=SMALL INTESTINE;			
RX MEDLINE=21085660; PubMed=11217851;			
RA Kawai J., Shindagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,			
RA Atakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,			
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,			

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT Functional annotation of a full-length mouse cDNA collection.;  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK008242; BAB2553.1; -;  
 DR MGD: MGI:1914185; 1700023B02R1K.  
 SQ SEQUENCE 277 AA; 31797 MW; 48EAD95516FBAC32 CRC64;

Query Match 29.8%; Score 51; DB 11; Length 277;  
 Best Local Similarity 46.7%; Pred. No. 13;  
 Matches 14; Conservative 2; Mismatches 8; Indels 6; Gaps 1;  
 Oy 1 SRTSRHSMETPTDINPAMVASRCIRPVGR 30  
 Db 200 SRSRHSPEKKGSDRN-----RGIRSR 223

RESULT 13  
 ID 09DAE5 PRELIMINARY; PRT; 286 AA.  
 AC 09DAE5;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE 1700023B02R1K PROTEIN.  
 GN 1700023B02R1K.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT Functional annotation of a full-length mouse cDNA collection.;  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK005900; BAB24307.1; -;  
 DR MGD: MGI:1914185; 1700023B02R1K.  
 SQ SEQUENCE 286 AA; 32818 MW; 068003C5E894827B CRC64;

Query Match 29.8%; Score 51; DB 11; Length 286;  
 Best Local Similarity 46.7%; Pred. No. 13;  
 Matches 14; Conservative 2; Mismatches 8; Indels 6; Gaps 1;  
 Oy 1 SRTSRHSMETPTDINPAMVASRCIRPVGR 30  
 Db 209 SRSRHSPEKKGSDRN-----RGIRSR 232

RESULT 14  
 ID 09DA19 PRELIMINARY; PRT; 450 AA.  
 AC 09DA19;  
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE 1700023B02R1K PROTEIN.  
 GN 1700023B02R1K.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OC NCBI\_TaxID=10090;  
 OX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=TESTIS;  
 RX MEDLINE=21085660; PubMed=11217851;  
 RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
 RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilming L.,  
 RA Wyszynski-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohetsuki S.,  
 RA Hayashizaki Y.,  
 RT Functional annotation of a full-length mouse cDNA collection.;  
 RL Nature 409:685-690(2001).  
 DR EMBL: AK006260; BAB24488.1; -;  
 DR MGD: MGI:1914185; 1700023B02R1K.  
 DR InterPro: IPR000345; CytC\_heme\_bind.  
 DR PROSITE: PS00190; CYTOCHROME\_C; UNKNOWN\_1.  
 SQ SEQUENCE 450 AA; 51852 MW; F32F11B6D64EAC CRC64;

Query Match 29.8%; Score 51; DB 11; Length 450;  
 Best Local Similarity 46.7%; Pred. No. 22;  
 Matches 14; Conservative 2; Mismatches 8; Indels 6; Gaps 1;  
 Oy 1 SRTSRHSMETPTDINPAMVASRCIRPVGR 30  
 Db 373 SRSRHSPEKKGSDRN-----RGIRSR 396

RESULT 15  
 ID 060687 PRELIMINARY; PRT; 465 AA.  
 AC 060687;  
 DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
 DT 01-AUG-1998 (TrEMBLrel. 07, Last sequence update)  
 DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
 DE SUSHI-REPEAT PROTEIN.  
 GN SRPUL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kurosawa H., Inukai T., Inaba T., Gol K., Chang K.-S., Sanjyo T.,  
 RL Submitted (Apr-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AF060567; AAC15765.1; -.  
 DR InterPro: IPR001128; Cyt\_P450.  
 DR InterPro: IPR003410; HVR.  
 DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
 DR Pfam: PF02494; HVR; 1.  
 DR Pfam: PF00084; sush1; 3.  
 DR SMART: SM00032; CCP; 3.  
 DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
 SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FE3EFB8 CRC64;

Query Match 29.8%; Score 51; DB 4; Length 465;  
 Best Local Similarity 50.0%; Pred. No. 23;  
 Matches 8; Conservative 1; Mismatches 7; Indels 0; Gaps 0;

QY 12 TPDINPAWYASRGIRP 27  
 DB 18 TPAVTPWTAGSGYYP 33

Search completed: September 13, 2002, 09:29:16  
 Job time: 1061 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:11:50 : Search time 80.21 seconds  
(without alignments)  
14.965 Million cell updates/sec

Title: US-09-446-543A-61  
Perfect score: 171  
Sequence: 1 SRTRHSMERTPDINPAWASRGIRPVGRF 31

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	171	100.0	87	1	PRRP_HUMAN
2	158	92.4	98	1	PRRP_BOVIN
3	159	87.1	83	1	PRRP_RAT
4	54	31.6	428	1	NER3_HUMAN
5	52	30.4	288	1	Y587_PASPU
6	50.5	29.5	1882	1	Y468_MYCPN
7	50	29.2	428	1	NER3_BOVIN
8	47.5	27.8	345	1	ARGC_BCHND
9	47.5	27.8	533	1	CNC_DROME
10	47	27.5	402	1	EX7L_STRCO
11	46.5	27.2	676	1	PLSC_HELPY
12	46.5	27.2	240	1	PLSC_HELPY
13	46.5	27.2	1693	1	POLN_HEVBU
14	46.5	27.2	1693	1	POLN_HEVBU
15	46.5	27.2	1693	1	POLN_HEVBU
16	46	26.9	342	1	Y762_METJA
17	46	26.9	347	1	Y576_METJA
18	46	26.9	383	1	CYCR_CHRVI
19	45.5	26.6	239	1	6PGL_SYNY3
20	45.5	26.6	407	1	K116_MYCTU
21	45	26.3	213	1	SRN2_YEAST
22	45	26.3	350	1	HMRO_DROME
23	45	26.3	510	1	YCGH_ECOLI
24	45	26.3	798	1	UNR_RAT
25	45	26.3	962	1	UNRA_METTH
26	45	26.3	973	1	UNRA_METTH
27	44.5	26.0	860	1	VG12_BPBO3
28	44	25.7	137	1	SMR2_RAT
29	44	25.7	364	1	YAIM_ECOLI
30	44	25.7	386	1	CRTY_AGRAL
31	44	25.7	476	1	YAAJ_ECOLI
32	44	25.7	481	1	POL_MLYRK
33	44	25.7	591	1	PIRG_HUMAN

34	44	25.7	719	1	NRPI_YEAST
35	44	25.7	843	1	POL_MLYRK
36	44	25.7	1087	1	PR16_HUMAN
37	44	25.7	1087	1	PR16_MOUSE
38	44	25.7	1196	1	POL_MLYRK
39	44	25.7	1196	1	POL_MLYRK
40	43.5	25.4	568	1	RS1_RICPR
41	43.5	25.4	572	1	GAG_IPHA
42	43.5	25.4	770	1	AVP3_ARATH
43	43.5	25.4	870	1	PLSR_XYLA
44	43.5	25.4	1316	1	RPOC_MYCLE
45	43	25.1	162	1	PHCA_CYACA

## ALIGNMENTS

RESULT ID	PRRP_HUMAN	STANDARD	PRT	87 AA
AC	P81277			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-releasing peptide PRRP20].			
DE	releasing peptide PRRP20].			
GN	PRP			
OS	Homo sapiens (human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=98268761; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kikada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RA	"A proactin-releasing peptide in the brain.";			
RT	Nature 393:272-276(1998).			
RL	[2]			
RN	TISSUE SPECIFICITY.			
RP	PubMed=10498338;			
RX	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kikada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;			
RA	"Tissue distribution of proactin-releasing peptide (PRP) and its receptor.";			
RT	Regul. Pept. 83:1-10(1999).			
RL	-1- FUNCTION: Stimulates proactin (PR) release and regulates the expression of proactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PR.			
CC	-1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
DR	EMBL: AB015419; BAA29027.1; -			
DR	HMM: 602663;			
KW	Hormone; Amidation; Signal.			
FT	SIGNAL			
FT	PEPTIDE			
FT	MOD. RES			
FT	SEQUENCE			
FT	BY SIMILARITY.			
FT	PROACTIN-RELEASING PEPTIDE PRRP31.			
FT	PROACTIN-RELEASING PEPTIDE PRRP20.			
FT	AMIDATION (G-54 PROVIDE AMIDE GROUP).			
FT	MOD. RES			
FT	SEQUENCE			

Query Match 100.0%; Score 171; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 9.8e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 31  
 |||  
 23 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 53

RESULT 2  
 PRRP\_BOVIN STANDARD; PRT; 98 AA.  
 ID PRRP\_BOVIN  
 AC P81264:  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].  
 DE PRH.  
 GN Bos taurus (Bovine).  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidea;  
 OC Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kurokawa T., Masuo Y., Asano T., Matsunoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC CC  
 CC EMBL: AB015417; BAA29025.1;  
 DR Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 KW SIGNAL: 1 22  
 FT PEPTIDE: 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE: 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD.RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 92.4%; Score 158; DB 1; Length 98;  
 Best Local Similarity 90.3%; Pred. No. 9.6e-17;  
 Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

1 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 31  
 |||  
 23 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 53

RESULT 3  
 PRRP\_RAT STANDARD; PRT; 83 AA.  
 ID PRRP\_RAT  
 AC P81278:  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing

hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-releasing peptide PrRP20].  
 DE PRH.  
 GN Rattus norvegicus (Rat).  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kurokawa T., Masuo Y., Asano T., Matsunoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 CC [2]  
 RN TISSUE SPECIFICITY.  
 RP PubMed=10498338;  
 RX Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RA "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor.";  
 RT Regul. Pept. 83:1-10(1999).  
 RL CC  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.  
 CC CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC CC  
 CC EMBL: AB015418; BAA29026.1;  
 DR Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 KW SIGNAL: 1 21  
 FT PEPTIDE: 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE: 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD.RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 83 AA; 9215 MW; DDC75A264EE84F29 CRC64;

Query Match 87.1%; Score 149; DB 1; Length 83;  
 Best Local Similarity 83.9%; Pred. No. 1.8e-15;  
 Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

1 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 31  
 |||  
 22 SRTHRSMEIRTPDINPAMTGTGIRPVGRF 52

RESULT 4  
 NER3\_HUMAN STANDARD; PRT; 428 AA.  
 ID NER3\_HUMAN  
 AC Q9U049; Q9NOE1;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Stalldase 3 (EC 3.2.1.18) (Membrane stalldase) (Ganglioside stalldase) (N-acetyl-alpha-neuraminidase 3).  
 DE NEU3.  
 GN Homo sapiens (Human).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

CC TISSUE-Brain; PubMed-10405317;  
RA MEDLINE-99353533; PubMed-10405317;  
RA Wada T., Yoshikawa Y., Tokuyama S., Kuwabara M., Akita H., Miyagi T.;  
RT "Cloning, expression, and chromosomal mapping of a human ganglioside  
RT sialidase";  
RL Biochem. Biophys. Res. Commun. 261:21-27(1999).  
RN (2)  
RP SEQUENCE FROM N.A.  
RC TISSUE-Skeletal muscle;  
RX PubMed-10861246;  
RA Monti G., Basel M.T., Papini N., Riboni M., Manzoni M., Venerando B.,  
RA Croci G., Preti A., Ballabio A., Tettamanzi G., Borsani G.;  
RT Identification and expression of NEU3, a novel human sialidase  
RT associated to the Plasma membrane";  
RL Biochem. J. 349:343-351(2000).  
CC -1- FUNCTION: Plays a role in modulating the ganglioside content of  
CC the lipid bilayer at the level of membrane-bound sialyl  
CC glycoconjugates.  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,  
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in  
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and  
CC synthetic substrates.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated.  
CC -1- TISSUE SPECIFICITY: Highly expressed in skeletal muscle, testis,  
CC adrenal gland and thymus; followed by pancreas, liver, heart and  
CC thymus. Weakly expressed in kidney, placenta, brain and lung.  
CC -1- MISCELLANEOUS: Optimum pH is 3.8.  
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -1- SIMILARITY: CONTAINS 3 BNR REPEATS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
CC EMBL: AB008185; BAA82611.1; -  
DR EMBL: Y18563; CAB96131.1; ALT\_INIT.  
DR MIM: 604617; -  
DR InterPro: IPR002860; BNR.  
DR Pfam: PF02012; BNR. 3.  
KW Hydrolase; Glycosidase; Membrane; Repeat.  
FT REPEAT 129 140 BNR 1.  
FT REPEAT 203 214 BNR 2.  
FT REPEAT 254 265 BNR 3.  
FT SITE 24 27 FRIP MOTIF.  
FT ACT\_SITE 25 25 BY SIMILARITY.  
FT ACT\_SITE 45 45 POTENTIAL.  
FT ACT\_SITE 50 50 POTENTIAL.  
FT ACT\_SITE 87 87 POTENTIAL.  
FT ACT\_SITE 225 225 POTENTIAL.  
FT ACT\_SITE 245 245 POTENTIAL.  
FT ACT\_SITE 340 340 BY SIMILARITY.  
FT ACT\_SITE 370 370 POTENTIAL.  
FT ACT\_SITE 387 387 POTENTIAL.  
SO SEQUENCE 428 AA; 48252 MW; 35D1DD9359A78C98 CRC64;

Query Match 31.6%; Score 54; DB 1; Length 428;  
Best Local Similarity 37.0%; Pred. No. 1.5;  
Matches 10; Conservative 6; Mismatches 11; Indels 0; Gaps 0;

OY 2 RTHSHMETPDPINPAWASRGIRPV 28  
Db 195 KTRPHSLMTYSDDLGVTHHGLIRPM 221

RESULT 5  
ID Y587\_PASMU STANDARD; PRT; 288 AA.  
AC Q9CNE6;

DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein PM0587.  
GN PM0587.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-PM70;  
RX MEDLINE-21145866; PubMed-11248100;  
RA May B.J., Zhang Q., Li L.L., Pauslan M.L., Whitlam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida PM70";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -1- SIMILARITY: BELONGS TO THE FRUCTOSAMINE KINASE FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: AE006094; AAK02671.1; -  
DR Hypothetical protein; Transferase; Kinase; Complete proteome.  
KM SEQUENCE 288 AA; 33778 MW; F4D2P6C26014D940 CRC64;

Query Match 30.4%; Score 52; DB 1; Length 288;  
Best Local Similarity 37.5%; Pred. No. 1.9;  
Matches 9; Conservative 6; Mismatches 9; Indels 0; Gaps 0;

OY 5 RSHSHMETPDPINPAWASRGIRPV 28  
Db 20 KHKEKHTGEMHFAWIIDGIOPV 43

RESULT 6  
ID Y468\_MYCPN STANDARD; PRT; 1882 AA.  
AC P75109; Q50317;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MG468 homolog (K05\_orf1882).  
GN MPN684 OR MP158.  
OS Mycoplasma pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Molluscites;  
OC Mycoplasmataceae; Mycoplasma.  
OX NCBI\_TaxID=2104;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 29342 / M129;  
RX MEDLINE-97105885; PubMed-8948633;  
RA Himmelreich R., Hilbert H., Plegens H., Pirkl E., Li B.-C.,  
RA Hermann R.;  
RT "Complete sequence analysis of the genome of the bacterium Mycoplasma  
RT pneumoniae";  
RL Nucleic Acids Res. 24:4420-4449(1996).  
RN [2]  
RP SEQUENCE OF 1-1848 FROM N.A.  
RC STRAIN-ATCC 29342 / M129;  
RX MEDLINE-96177562; PubMed-8604303;  
RA Hilbert H., Himmelreich R., Plegens H., Hermann R.;  
RT "Sequence analysis of 56 kb from the genome of the bacterium  
RT Mycoplasma pneumoniae comprising the dnaA region, the atp operon and a  
RT cluster of ribosomal protein genes";  
RL Nucleic Acids Res. 24:628-639(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: SOME, TO MG064.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AE000017; AAB95806.1; -  
DR EMBL: U34816; AAC43650.1; -  
KW Hypothetical protein; Transmembrane; Complete proteome.  
FT TRANSMEM 16 36 POTENTIAL.  
FT TRANSMEM 987 1007 POTENTIAL.  
FT TRANSMEM 1037 1057 POTENTIAL.  
FT TRANSMEM 1080 1100 POTENTIAL.  
FT TRANSMEM 1154 1174 POTENTIAL.  
FT TRANSMEM 1759 1779 POTENTIAL.  
FT TRANSMEM 1807 1827 POTENTIAL.  
FT TRANSMEM 1828 1848 POTENTIAL.  
FT TRANSMEM 1851 1871 POTENTIAL.  
SQ SEQUENCE 1882 AA; 209442 MW; 03CPAD99A7120ED CRC64;

Query Match 29.5%; Score 50.5; DB 1; Length 1882;  
Best Local Similarity 32.3%; Pred. No. 25;  
Matches 10; Conservative 6; Mismatches 12; Indels 3; Gaps 1;

OY 1 SRTHSMETRPDINPAWYASRGIRPV 31  
DB 1188 NNFNRYRLMLQPTQSGWYA---IQPSRF 1215

RESULT 7  
NER3\_BOVIN STANDARD; PRT; 428 AA.  
AC 097859;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Stalidase 3 (EC 3.2.1.18) (Membrane stalidase) (Ganglioside stalidase)  
DE (N-acetyl-alpha-neuraminidase 3).  
CN NEU3.  
OS Bos taurus (bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., PARTIAL SEQUENCE, AND CHARACTERIZATION.  
RC TISSUE=Brain;  
RX MEDLINE=99143165; PubMed=9988745;  
RA Miyagi T., Wada T., Iwamatsu A., Hata K., Yoshikawa Y., Tokuyama S.,  
RA Sawada M.;  
RT "Molecular cloning and characterization of a plasma membrane-  
RT associated stalidase specific for gangliosides.";  
RL J. Biol. Chem. 274:5004-5011(1999).  
CC -1- FUNCTION: Plays a role in modulating the ganglioside content of  
CC the lipid bilayer at the level of membrane-bound sialyl  
CC glycoconjugates.  
CC -1- CATALYTIC ACTIVITY: Hydrolysis of alpha-(2->3)-, alpha-(2->6)-,  
CC alpha-(2->8)-glycosidic linkages of terminal sialic residues in  
CC oligosaccharides, glycoproteins, glycolipids, colominic acid and  
CC synthetic substrates.  
CC -1- SUBCELLULAR LOCATION: Membrane-associated (By similarity).  
CC -1- TISSUE SPECIFICITY: Expressed in brain.  
CC -1- SIMILARITY: BELONGS TO FAMILY 33 OF GLYCOSYL HYDROLASES.  
CC -1- SIMILARITY: CONTAINS 3 BNR REPEATS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AB008184; BAA75071.1; -  
DR InterPro: IPR002860; BNR.  
KW Hydrolase; Glycosidase; Membrane; Repeat.  
FT REPEAT 129 140 BNR 1.  
FT REPEAT 203 214 BNR 2.  
FT REPEAT 254 265 BNR 3.  
FT SITE 24 27 FRIP MOTIF.  
FT ACT\_SITE 25 25 BY SIMILARITY.  
FT ACT\_SITE 45 45 POTENTIAL.  
FT ACT\_SITE 50 50 POTENTIAL.  
FT ACT\_SITE 87 87 POTENTIAL.  
FT ACT\_SITE 225 225 POTENTIAL.  
FT ACT\_SITE 245 245 POTENTIAL.  
FT ACT\_SITE 341 341 BY SIMILARITY.  
FT ACT\_SITE 371 371 POTENTIAL.  
FT ACT\_SITE 388 388 POTENTIAL.  
SQ SEQUENCE 428 AA; 47916 MW; 418B34EF3245ABF21 CRC64;

Query Match 29.2%; Score 50; DB 1; Length 428;  
Best Local Similarity 33.3%; Pred. No. 5.9;  
Matches 9; Conservative 6; Mismatches 12; Indels 0; Gaps 0;

OY 2 RTHRSMETRPDINPAWYASRGIRPV 28  
DB 195 RARPHSLMTYSDLCATYHGRILKPW 221

RESULT 8  
ARGC\_BACHD STANDARD; PRT; 345 AA.  
AC 09K8V2;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-  
DE acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).  
CN ARGC OR B2990.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hirata C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis."  
RL Nucleic Acids Res. 28:4317-4331(2000).  
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+) +  
CC + phosphate = N-acetyl-5-glutamyl phosphate + NADPH.  
CC -1- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.  
CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AP001517; BAB06619.1; -  
DR InterPro: IPR00706; AGPR act. site.  
DR InterPro: IPR00534; Semialdh\_ch.

DR Pfam: PF01118; Semaldehyde\_dh; 1.  
 DR Pfam: PF02774; Semaldehyde\_dhc; 1.  
 DR ProDom: PD003765; AGPR\_act\_site; 1.  
 DR PROSITE: PS01224; ARGC; 1.  
 KM Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.  
 FT ACT SITE 149 149 BY SIMILARITY.  
 SQ SEQUENCE 345 AA; 38188 MW; 3E9F45DD09FC68EA CRC64;

Query Match 27.8%; Score 47.5; DB 1; Length 345;  
 Best Local Similarity 45.5%; Pred. No. 11;  
 Matches 10; Conservative 2; Mismatches 9; Indels 1; Gaps 1;

QY 8 MEIRTPDINPAMYASRGIRPVG 29  
 : | | | | | | | | | |  
 Db 101 LRINEPDVYEAWY-KROAPVG 121

## RESULT 9

CNC\_DROME CNC\_DROME STANDARD; PRT; 533 AA.  
 AC P20482;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Segmentation protein cap'n collar.  
 GN CNC.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE:92001535; PubMed-1911393;  
 RA Mohler J., Vanl K., Leung S., Epstein A.;  
 RT "Segmentally restricted, cephalic expression of a leucine zipper gene during Drosophila embryogenesis.";  
 RL Mech. Dev. 34:3-9(1991).  
 CC -1- FUNCTION: PLAYS A ROLE IN CEPHALIC PATTERNING. PROBABLE SUBUNIT OF A HETERODIMERIC REGULATORY PROTEIN INVOLVED IN THE CONTROL OF HEAD MORPHOGENESIS.  
 CC -1- DEVELOPMENTAL STAGE: LOCALIZED TO THE MANDIBULAR SEGMENT AND THE HYPOPHARYNGEAL AND LABRAL PRIMORDIA FIRST DETECTABLE IN LATE BLASTODERM STAGES.  
 CC -1- SIMILARITY: BELONGS TO THE BZIP FAMILY. CNC SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M37495; AAB59246.1; -  
 DR PIR: A33111; A33111.  
 DR HSSP: P34707; 1SKN.  
 DR TRANSFAC: T01998; -  
 DR FLYBASE: FBgn0000338; CNC.  
 DR InterPro: IPR001871; bZIP.  
 DR Pfam: PF00170; bZIP; 1.  
 DR SMART: SM00338; BRLZ; 1.  
 DR PROSITE: PS00036; bZIP\_BASIC; 1.  
 KM Transcription regulation; Activator; DNA-binding; Nuclear protein;  
 KM Developmental protein.  
 FT DNA\_BIND 351 369  
 FT DOMAIN 387 408 BASIC MOTIF.  
 FT DOMAIN 497 529 LECICINE-ZIPPER.  
 FT CONFLICT 349 349 GLN-RICH.  
 FT CONFLICT 349 349 I -> L (IN REF. 1; AAB59246).  
 SQ SEQUENCE 533 AA; 56948 MW; EADPF9A5D6CA5C5F CRC64;

Query Match 27.8%; Score 47.5; DB 1; Length 533;  
 Best Local Similarity 37.5%; Pred. No. 18;  
 Matches 9; Conservative 5; Mismatches 9; Indels 1; Gaps 1;

QY 4 HRHSME-IRTPDINPAMYASRGIR 26  
 : | | | | | | | | | |  
 Db 409 HRHFQYLRDEEGNCPADYSIQ 432

## RESULT 10

EX7L\_STRCO EX7L\_STRCO STANDARD; PRT; 402 AA.  
 AC G9FBM3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
 DE (Exonuclease VII large subunit).  
 GN XSEA OR SCK7.29C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: BIDIRECTIONALLY DEGRADABLE SINGLE-STRANDED DNA INTO LARGE ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' - to 3' - or 3' - to 5' -direction to yield 5'-phosphomononucleotides.  
 CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AL391754; CAC05901.1; -  
 DR InterPro: IPR003753; Exonuc\_VII\_L.  
 DR InterPro: IPR002309; tRNA-synt\_2.  
 DR Pfam: PF02601; Exonuc\_VII\_L; 1.  
 DR Pfam: PF01336; tRNA\_anti\_1.  
 KM Hydrolase; Nuclease; Exonuclease.  
 SQ SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 27.5%; Score 47; DB 1; Length 402;  
 Best Local Similarity 47.6%; Pred. No. 15;  
 Matches 10; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 17 PAMYASRG-----IRPVG 29  
 : | | | | | | | | | |  
 Db 89 PEWYAPRGQLSLRAEIKIPVG 109

## RESULT 11

EX7L\_HUMAN EX7L\_HUMAN STANDARD; PRT; 676 AA.  
 AC G92935;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Exostosin-like I (Exostosin-L) (Multiple exostosin-like protein).



```

CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.1sb-slb.ch/announce/
CC or send an email to license@1sb-slb.ch).
CC -----
DR EMBL: D10330.1; BAA01172.1; -.
DR MEROPS: C41.001; -.
DR InterPro: IPR002589; Altp.
DR InterPro: IPR001788; RNA_dep_RNAPol2.
DR InterPro: IPR002588; V_methyltransf.
DR InterPro: IPR000606; Viral_helicase1.
DR Pfam: PF01661; Alpp. 1.
DR Pfam: PF00978; RNA_dep_RNAPol2; 1.
DR Pfam: PF01443; Viral_helicase1; 1.
DR Pfam: PF01660; Vmethyltransf; 1.
DR SMART: SM00506; Alpp; 1.
DR Polyprotein; Transferase; RNA-directed RNA polymerase; Helicase;
KW ATP-binding.
FT NP_BIND 975 982 ATP (POTENTIAL).
SQ SEQUENCE 1693 AA; 185215 MW; AAB4C9140A7E21EA CRC64;

Query Match 27.28; Score 46.5; DB 1; Length 1693;
Best Local Similarity 37.08; Pred. No. 87;
Matches 10; Conservative 5; Mismatches 11; Indels 1; Gaps 1;

Db 2 RTHRSMERITPDINPAMY-ASRGIRP 27
1 11 1 1 1 1 1 1 1 1 1 1
Db 904 RNHRGDELRYPELAAHFENRPTRP 930

RESULT 15
Y468_MYCGE STANDARD; PRT; 1783 AA.
AC 049460;
DT 01-NOV-1997 (Rel. 35, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein MG468.
GN MG468.
OS Mycoplasma genitalium.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Mollicutes;
OC Mycoplasmataceae; Mycoplasma.
OX NCBI_TaxID=2097;
RN 1;
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=96026346; PubMed=7569993;
RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RA Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RA Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RA "The minimal gene complement of Mycoplasma genitalium.";
RT Science 270:397-403(1995).
RL 12;
RN 12;
RP REVISIONS.
RP Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,
RP Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,
RP Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhrmann J.L.,
RP Nguyen D.T., Uterback T.R., Saudek D.M., Phillips C.A., Merrick J.M.,
RP Tomb J.F., Dougherty B.A., Bott K.F., Hu P.-C., Lucier T.S.,
RP Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;
RP Submitted (OCT-1998) to the EMBL/GenBank/DBJ databases.
RP [3]
RP SEQUENCE OF 879-985 FROM N.A.
RC STRAIN=ATCC 33530 / G-37;
RX MEDLINE=94075230; PubMed=8253680;
RA Peterson S.N., Hu P.-C., Bott K.F., Hutchison C.A. III;
RT "A survey of the Mycoplasma genitalium genome by using random
RT sequencing.";

```





GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:11:05 ; Search time 172.41 Seconds  
(without alignments)  
17.277 Million cell updates/sec

Title: US-09-446-543a-61

Perfect score: 171

Sequence: 1 SRRHSMETRPDINPAWYASRGIRPVGRF 31

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71:\*

1: PIR1:\*

2: PIR2:\*

3: PIR3:\*

4: PIR4:\*

## SUMMARIES

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

Result No.	Score	Query Match	Length	DB ID	Description
1	149	87.1	83	2 JC7607	prolactin-releasin
2	56.5	33.0	664	2 F83376	conserved hypothet
3	54	31.6	790	2 T47959	hypothetical prote
4	53.5	31.3	314	2 B70569	hypothetical prote
5	53	31.0	637	2 D87559	sensory box histid
6	52	30.4	785	2 F69099	sensory transduct
7	50.5	29.5	1882	2 S73484	hypothetical prote
8	50	29.2	123	2 S77900	hypothetical prote
9	50	29.2	128	2 S76955	hypothetical prote
10	49.5	28.9	176	2 S67150	hypothetical prote
11	49	28.7	72	2 E91002	probable regulator
12	48.5	28.4	303	2 AH2016	hypothetical prote
13	48.5	28.4	1501	2 T45623	hypothetical prote
14	48	28.1	118	2 AC3169	hypothetical prote
15	48	28.1	220	2 C83292	probable glutathio
16	48	28.1	335	2 S70671	lipopolysaccharide
17	48	28.1	348	2 T21648	hypothetical prote
18	48	28.1	455	2 D70885	probable aldC prot
19	48	28.1	1084	2 T33759	hypothetical prote
20	47.5	27.8	345	2 D84012	N-acetylglutamate
21	47.5	27.8	393	2 AB2664	conserved hypothet
22	47.5	27.8	401	2 A97446	hypothetical prote
23	47.5	27.8	503	2 A82193	Sun/nucleolar prote
24	47.5	27.8	533	2 A33111	segmentation prote
25	47	27.5	159	2 G82669	ubiquitome biosynt
26	47	27.5	215	2 B87577	glutathione S-tran
27	47	27.5	284	2 F71015	hypothetical prote
28	47	27.5	333	2 H82852	hydroxybenzoate oc
29	47	27.5	501	2 T48336	hypothetical prote

30	47	27.5	938	2 C84480	hypothetical prote
31	47	27.5	4589	2 T14914	dynamin beta heavy
32	46.5	27.2	240	2 D64688	probable 1-acylgly
33	46.5	27.2	779	2 T49717	related to BCS1 pr
34	46.5	27.2	957	2 AB4089	hypothetical prote
35	46.5	27.2	1495	2 T31434	denin-180 - rat
36	46.5	27.2	1607	2 T13250	hypothetical prote
37	46.5	27.2	1693	2 MNMWE	genome polyprotein
38	46.5	27.2	4957	2 T03455	ALR protein - huma
39	46.5	27.2	5262	2 T03454	ALR protein - huma
40	46	26.9	256	2 F70812	probable 1pqr prot
41	46	26.9	342	2 B64395	malic acid transpo
42	46	26.9	347	2 H64371	malic acid transpo
43	46	26.9	419	2 AH3166	hypothetical prote
44	46	26.9	688	2 AT2516	hypothetical prote
45	46	26.9	698	2 T39050	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R.Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, B.; Blochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YMM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 87.1%; Score 149; DB 2; Length 83;  
Best Local Similarity 83.9%; Pred. No. 8.5e-15;  
Matches 26; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 SRRHSMETRPDINPAWYASRGIRPVGRF 31  
DB 22 SRAHSMETRPDINPAWYTGIRPVGRF 52

RESULT 2  
F83376  
conserved hypothetical protein PA2151 [imported] - Pseudomonas aeruginosa (strain PAO  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83376  
R.Stover, C.K.; Pham, X.Q.; Edwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kass, A.; Lardig, K.; L  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PAO1, an opportunistic pa  
A:Reference number: AB2950; MUID:20437337  
A:Accession: F83376  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-664 <STO>  
A:Cross-references: GB:AE004642; GB:AE004091; NID:g9948163; PIDN:AA605539.1; GSPDB:GN  
A:Experimental source: strain PAO1  
C:Genetics:  
A:Gene: PA2151

Query Match 33.0%; Score 56.5; DB 2; Length 664;  
Best Local Similarity 45.8%; Pred. No. 3.4;  
Matches 11; Conservative 4; Mismatches 8; Indels 1; Gaps 1;

OY 4 HRSMEIRTPDINPAWYASRGIRP 27  
DB 470 YRPNFVNTPDINP-WFLORSGRP 492  
| : : | | | | | : | : | |

RESULT 3  
T47959  
hypothetical protein F15G16\_60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Quett  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: 224480  
A:Accession: T47959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEH>  
A:Cross-References: EMBL:AL132959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 39/1; 678/2; 698/3; 773/2  
A>Note: F15G16\_60

Query Match 31.6%; Score 54; DB 2; Length 790;  
Best Local Similarity 52.6%; Pred. No. 9.7;  
Matches 10; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 13 PDINPAWYASRGIRPGRF 31  
DB 366 PPHNRTYGSRLQPHGRW 384  
| | | | | | | | | |

RESULT 4  
B70569  
hypothetical protein RV3485c - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70569  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.  
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.  
; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
; Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987  
A:Accession: B70569  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-314 <COI>  
A:Cross-References: GB:Z95390; GB:AL123456; NID:93261766; PIDN:CAB08708.1; PID:92104408  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV3485c  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology  
F:46-227/Domain: short-chain alcohol dehydrogenase homology <SADR>

Query Match 31.3%; Score 53.5; DB 2; Length 314;  
Best Local Similarity 27.5%; Pred. No. 4.2;  
Matches 11; Conservative 5; Mismatches 11; Indels 13; Gaps 1;

OY 1 SRTHR-----HSMEIRTPDINPAWYASRGIRP 27  
DB 187 SNTHRWFGAYGVTKSAVDHMKMLADELGPSSVVRVNSIRP 226  
| | | | | : : | : | | | |

RESULT 5  
D87559  
sensory box histidine kinase/response regulator [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: D87559  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, R.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.D.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete genome sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: D87559  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-637 <STO>  
A:Cross-References: GB:AE005673; NID:g13424056; PIDN:AAK24472.1; GSFDB:GN00148  
C:Genetics:  
A:Gene: CC2501

Query Match 31.0%; Score 53; DB 2; Length 637;  
Best Local Similarity 48.0%; Pred. No. 11;  
Matches 12; Conservative 4; Mismatches 9; Indels 0; Gaps 0;

OY 4 HRSMEIRTPDINPAWYASRGIRPV 28  
DB 22 HRSDDLRSPAINPAIRVRIRAV 46  
| | : : | | | | | : | |

RESULT 6  
F69099  
sensory transduction histidine kinase - Methanobacterium thermoautotrophicum (strain  
C:Species: Methanobacterium thermoautotrophicum  
C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 22-Oct-1999  
C:Accession: F69099  
R:Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T  
; Qiu, D.; Spadafora, R.; Vitale, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jivani, L  
; S.; Church, G.M.; Daniels, C.J.; Mao, J.; Rice, P.; Noelling, J.; Reeve, J.N.  
J. Bacteriol. 179, 7135-7155, 1997  
A:Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: fu  
A:Reference number: A69000; MUID:98037514  
A:Accession: F69099  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-785 <MTH>  
A:Cross-References: GB:AE00805; GB:AE00666; NID:g2621213; PIDN:AB84680.1; PID:g262  
A:Experimental source: strain Delta H  
C:Genetics:  
A:Gene: MTH174

Query Match 30.4%; Score 52; DB 2; Length 785;  
Best Local Similarity 46.4%; Pred. No. 19;  
Matches 13; Conservative 1; Mismatches 12; Indels 2; Gaps 1;

OY 4 HRSMEIRTPDINPAW--YASRGIRPVG 29  
DB 412 HRATFRIRRPDNNYRWEYVDRPIRTDG 439  
| : : | | | | | | | | | |

RESULT 7  
S73484  
hypothetical protein K05\_0r1882 - Mycoplasma pneumoniae (strain ATCC 29342)  
C:Species: Mycoplasma pneumoniae  
A:Variety: ATCC 29342  
C:Date: 27-Feb-1997 #sequence\_revision 25-Apr-1997 #text\_change 07-Dec-1999  
C:Accession: S73484; S62840  
R:Himmelreich, R.; Hilbert, H.; Plagens, H.; Plickl, E.; Li, B.C.; Herrmann, R.  
Nucleic Acids Res. 24, 4420-4449, 1996  
A:Title: Complete sequence analysis of the genome of the bacterium Mycoplasma pneumon  
A:Reference number: S73327; MUID:97105885



hypothetical protein all1686 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH2016  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 6, 205-213, 2001  
A:Title: Complete genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena PCC 7120  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2016  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-303 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA078052.1; PID:g17135506; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: all1686

Query Match 28.4%; Score 48.5; DB 2; Length 303;  
Best Local Similarity 48.4%; Pred. No. 22;  
Matches 15; Conservative 1; Mismatches 10; Indels 5; Gaps 2;

OY 4 HRRHMEIRTPDINPAMY----ASRGIRPVGR 30  
DB 226 HEHSTE-RTRAIDGTYTLCGAGANRPVGR 255

RESULT 13  
T45623  
hypothetical protein F13624.180 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 04-Feb-2000  
C:Accession: T45623  
R:Bevan, M.; Van Der Schueren, J.; Chuang, Y.J.; Voet, M.; Robben, J.; Volckaert, G.; Bae submitted to the Protein Sequence Database, December 1999  
A:Reference number: 223009  
A:Accession: T45623  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1501 <BEV>  
A:Cross-references: EMBL:AL133421  
A:Experimental source: cultivar Columbia; BAC clone F13624  
C:Genetics:  
A:Map position: 5  
A:Insertions: 64/1; 739/3; 785/2; 1302/2; 1318/3; 1399/2; 1434/2  
A:Note: F13624.180

Query Match 28.4%; Score 48.5; DB 2; Length 1501;  
Best Local Similarity 29.5%; Pred. No. 1.3e+02;  
Matches 13; Conservative 7; Mismatches 9; Indels 15; Gaps 3;

OY 3 THRHSMER-----TPDINPAMYASRG-----IRPV---GRF 31  
DB 1169 TKNHSAVGDHQQISPOKAPSFYSOYGTFKNGLVQPVNDTGRF 1212

RESULT 14  
AC3169  
hypothetical protein Atus077 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AC3169  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm, A.; Sterner, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193

A:Accession: AC3169  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-118 <KUR>  
A:Cross-references: GB:AE008687; PIDN:NAL45769.1; PID:g17743503; GSPDB:GN00188  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atus077  
A:Genome: plasmid

Query Match 28.1%; Score 48; DB 2; Length 118;  
Best Local Similarity 34.4%; Pred. No. 9.3;  
Matches 11; Conservative 5; Mismatches 12; Indels 4; Gaps 1;

OY 2 RTHRSMERITPDINPAMYASRGIRPVGR 29  
DB 19 KLHKHNAARRRDPDIGARVQPATAPHRGLAGIG 50

RESULT 15  
C83292  
Probable glutathione S-transferase PA2821 [imported] - Pseudomonas aeruginosa (strain C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001  
C:Accession: C83292  
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lathig, K.; L.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen  
A:Reference number: AB2950; MUID:20437337  
A:Accession: C83292  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-220 <STO>  
A:Cross-references: GB:AE004709; GB:AE004091; NID:g9948904; PIDN:AAG06209.1; GSPDB:GN001821  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA2821  
A:Superfamily: plaiace glutathione transferase

Query Match 28.1%; Score 48; DB 2; Length 220;  
Best Local Similarity 44.0%; Pred. No. 18;  
Matches 11; Conservative 3; Mismatches 9; Indels 2; Gaps 1;

OY 6 HSMETRTPDINPAMYASRGIRPVGR 30  
DB 28 YOLEAIAPPGOPAWY--REISPLGR 50

Search completed: September 13, 2002, 09:23:57  
Job time: 772 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:10:35 : Search time 136.62 Seconds  
(without alignments)  
5.542 Million cell updates/sec

Title: US-09-446-543A-61  
Sequence: 1 SRTHSHSEIRTPDINPAWMSGIRPYGRF 31

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/CCTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfilset1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	171	100.0	31	3	US-09-105-678A-9
2	171	100.0	31	3	US-09-105-678A-43
3	171	100.0	31	4	US-08-776-971-61
4	171	100.0	31	4	US-09-421-208-9
5	171	100.0	31	4	US-09-421-208-43
6	171	100.0	32	3	US-09-105-678A-44
7	171	100.0	32	4	US-08-776-971-62
8	171	100.0	32	4	US-09-421-208-44
9	171	100.0	33	3	US-09-105-678A-45
10	171	100.0	33	4	US-08-776-971-63
11	171	100.0	33	4	US-09-421-208-45
12	171	100.0	87	4	US-08-776-971-59
13	171	100.0	87	4	US-08-776-971-135
14	171	100.0	87	4	US-08-776-971-138
15	158	92.4	31	3	US-09-105-678A-7
16	158	92.4	31	3	US-09-105-678A-31
17	158	92.4	31	4	US-08-776-971-5
18	158	92.4	31	4	US-08-776-971-97
19	158	92.4	31	4	US-09-421-208-7
20	158	92.4	31	4	US-09-421-208-31
21	158	92.4	32	4	US-09-105-678A-32
22	158	92.4	32	4	US-08-776-971-6
23	158	92.4	32	4	US-09-421-208-32
24	158	92.4	33	4	US-09-105-678A-33
25	158	92.4	33	4	US-08-776-971-7
26	158	92.4	33	4	US-09-421-208-33
27	158	92.4	98	4	US-08-776-971-1

28	158	92.4	98	4	US-08-776-971-44	Sequence 44, Appl
29	158	92.4	98	4	US-08-776-971-122	Sequence 122, App
30	158	92.4	98	4	US-08-776-971-131	Sequence 131, App
31	158	92.4	98	4	US-08-776-971-136	Sequence 136, App
32	154	90.1	98	4	US-08-776-971-115	Sequence 115, App
33	154	90.1	98	4	US-08-776-971-117	Sequence 117, App
34	149	87.1	31	3	US-09-105-678A-8	Sequence 8, Appl1
35	149	87.1	31	3	US-09-105-678A-37	Sequence 37, Appl1
36	149	87.1	31	4	US-09-172-353-4	Sequence 47, Appl
37	149	87.1	31	4	US-08-776-971-47	Sequence 8, Appl1
38	149	87.1	31	4	US-09-421-208-8	Sequence 37, Appl
39	149	87.1	31	4	US-09-421-208-37	Sequence 38, Appl
40	149	87.1	32	3	US-09-105-678A-38	Sequence 48, Appl
41	149	87.1	32	4	US-08-776-971-48	Sequence 38, Appl
42	149	87.1	32	4	US-09-421-208-38	Sequence 39, Appl
43	149	87.1	33	3	US-09-105-678A-39	Sequence 49, Appl
44	149	87.1	33	4	US-08-776-971-49	Sequence 39, Appl
45	149	87.1	33	4	US-09-421-208-39	

#### ALIGNMENTS

RESULT 1  
US-09-105-678A-9  
Sequence 9, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-9

Query Match 100.0% Score 171; DB 3; Length 31;  
Best Local Similarity 100.0% Pred. No. 6.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDPINPAMYASRGIRPVGRF 31  
DB 1 SRTHRSMETRPDPINPAMYASRGIRPVGRF 31

## RESULT 2

US-09-105-678A-43  
; Sequence 43, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-43

Query Match 100.0%; Score 171; DB 3; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6,1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDPINPAMYASRGIRPVGRF 31  
DB 1 SRTHRSMETRPDPINPAMYASRGIRPVGRF 31

## RESULT 3

US-08-776-971-61  
; Sequence 61, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; APPLICANT: Habata, Yugo  
; APPLICANT: Kawamata, Yuji  
; APPLICANT: Hosoya, Masaki  
; APPLICANT: Fujii, Ryo  
; APPLICANT: Fukusumi, Shoji  
; APPLICANT: Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 100.0%; Score 171; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6,1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDPINPAMYASRGIRPVGRF 31  
DB 1 SRTHRSMETRPDPINPAMYASRGIRPVGRF 31

## RESULT 4

US-09-421-208-9  
; Sequence 9, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 100.0%; Score 171; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31  
DB 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31

RESULT 5  
US-09-421-208-43  
Sequence 43, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-43

Query Match 100.0%; Score 171; DB 4; Length 31;  
Best Local Similarity 100.0%; Pred. No. 6.1e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31  
DB 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31

RESULT 6  
US-09-105-678A-44  
Sequence 44, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 44:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 32 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-44

Query Match 100.0%; Score 171; DB 3; Length 32;  
Best Local Similarity 100.0%; Pred. No. 6.3e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETPTDINPAMYASRGIRPVGRF 31  
 DB 1 SRTHRSMETPTDINPAMYASRGIRPVGRF 31

RESULT 7  
 US-08-776-971-62  
 ; Sequence 62, Application US/08776971B  
 ; Patent No. 6228984  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hinuma, Shuji  
 ; Habata, Yugo  
 ; Kawamata, Yuji  
 ; Hosoya, Masaki  
 ; Fujii, Ryo  
 ; Fukusumi, Shoji  
 ; Kitada, Chieko  
 ; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
 ; NUMBER OF SEQUENCES: 140  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Diskette  
 ; COMPUTER: IBM compatible  
 ; OPERATING SYSTEM: DOS  
 ; SOFTWARE: FASTSEQ for Windows Version 2.0  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/08/776,971B  
 ; FILING DATE: 06-Feb-1997  
 ; CLASSIFICATION: <unknown>  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: PCT/JP96/03821  
 ; FILING DATE: 28-DEC-1996  
 ; APPLICATION NUMBER: JP 7/343371  
 ; FILING DATE: 28-DEC-1995  
 ; APPLICATION NUMBER: JP 8/59419  
 ; FILING DATE: 15-MAR-1996  
 ; APPLICATION NUMBER: JP 8/211805  
 ; FILING DATE: 12-AUG-1996  
 ; APPLICATION NUMBER: JP 8/246573  
 ; FILING DATE: 18-SEP-1996  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Conlin, David G.  
 ; REGISTRATION NUMBER: 27,026  
 ; REFERENCE/DOCKET NUMBER: 47176  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-523-3400  
 ; TELEFAX: 617-523-6440  
 ; INFORMATION FOR SEQ ID NO: 62:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 32 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS: single  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: protein  
 ; FRAGMENT TYPE: internal  
 ; SEQUENCE DESCRIPTION: SEQ ID NO: 62:  
 ; US-08-776-971-62

Query Match 100.0%; Score 171; DB 4; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 6,3e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETPTDINPAMYASRGIRPVGRF 31  
 DB 1 SRTHRSMETPTDINPAMYASRGIRPVGRF 31

RESULT 8  
 US-09-421-208-44  
 ; Sequence 44, Application US/09421208  
 ; Patent No. 628561  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Suenaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street  
 ; CITY: Boston  
 ; STATE: MA  
 ; COUNTRY: USA  
 ; ZIP: 02109  
 ; COMPUTER READABLE FORM:  
 ; MEDIUM TYPE: Floppy disk  
 ; COMPUTER: IBM PC compatible  
 ; OPERATING SYSTEM: PC-DOS/MS-DOS  
 ; SOFTWARE: Patentin Release #1.0, Version #1.30  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/09/421,208  
 ; FILING DATE:  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: US 09/105,678  
 ; FILING DATE: 26-JUN-1998  
 ; APPLICATION NUMBER: JP 172118/1997  
 ; FILING DATE: 27-JUN-1997  
 ; ATTORNEY/AGENT INFORMATION:  
 ; NAME: Conlin, David G.  
 ; REGISTRATION NUMBER: 27,026  
 ; REFERENCE/DOCKET NUMBER: 48466-342  
 ; TELECOMMUNICATION INFORMATION:  
 ; TELEPHONE: 617-523-3400  
 ; TELEFAX: 617-523-6440  
 ; INFORMATION FOR SEQ ID NO: 44:  
 ; SEQUENCE CHARACTERISTICS:  
 ; LENGTH: 32 amino acids  
 ; TYPE: amino acid  
 ; STRANDEDNESS:  
 ; TOPOLOGY: linear  
 ; MOLECULE TYPE: peptide  
 ; US-09-421-208-44

Query Match 100.0%; Score 171; DB 4; Length 32;  
 Best Local Similarity 100.0%; Pred. No. 6,3e-19;  
 Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 SRTHRSMETPTDINPAMYASRGIRPVGRF 31  
 DB 1 SRTHRSMETPTDINPAMYASRGIRPVGRF 31

RESULT 9  
 US-09-105-678A-45  
 ; Sequence 45, Application US/09105678A  
 ; Patent No. 6103882  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Suenaga, Masato  
 ; APPLICANT: Moriya, Takeo  
 ; APPLICANT: Tanaka, Yoko  
 ; APPLICANT: Nishimura, Osamu  
 ; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
 ; NUMBER OF SEQUENCES: 52  
 ; CORRESPONDENCE ADDRESS:  
 ; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
 ; STREET: 130 Water Street



CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-45

Query Match 100.0%; Score 171; DB 3; Length 33;  
Best Local Similarity 100.0%; Pred. No. 6,6e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRTHRHSMETRPDINPAMYASRGIRPVGRF 31  
Db 1 SRTHRHSMETRPDINPAMYASRGIRPVGRF 31

RESULT 10  
US-08-776-971-63  
Sequence 63, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Tujl  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 63:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 63:  
US-08-776-971-63

Query Match 100.0%; Score 171; DB 4; Length 33;  
Best Local Similarity 100.0%; Pred. No. 6,6e-19;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 SRTHRHSMETRPDINPAMYASRGIRPVGRF 31  
Db 1 SRTHRHSMETRPDINPAMYASRGIRPVGRF 31

RESULT 11  
US-09-421-208-45  
Sequence 45, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 45:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 33 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-45

Query Match 100.0%; Score 171; DB 4; Length 33;  
Best Local Similarity 100.0%; Pred. No. 6.6e-15;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31  
DB 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31

RESULT 12  
US-08-776-971-59  
Sequence 59, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji1, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 59:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid

STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 59:  
US-08-776-971-59

Query Match 100.0%; Score 171; DB 4; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMETRPDINPAMYASRGIRPVGRF 31  
DB 23 SRTHRSMETRPDINPAMYASRGIRPVGRF 53

RESULT 13  
US-08-776-971-135  
Sequence 135, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji1, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 135:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal

SEQUENCE DESCRIPTION: SEQ ID NO: 135:  
US-08-776-971-135

Query Match 100.0%; Score 171; DB 4; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 31  
DB 23 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 53

## RESULT 14

US-08-776-971-138  
Sequence 138, Application US/08776971B  
Patent No. 6228984

## GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

## ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

## INFORMATION FOR SEQ ID NO: 138:

SEQUENCE CHARACTERISTICS:  
LENGTH: 87 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 138:

US-08-776-971-138

Query Match 100.0%; Score 171; DB 4; Length 87;  
Best Local Similarity 100.0%; Pred. No. 2e-18;  
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 31  
DB 23 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 53

## RESULT 15

US-09-105-678A-7  
Sequence 7, Application US/09105678A  
Patent No. 6103882

## GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
Moriya, Takeo  
Tanaka, Yoko  
Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

## COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998

## PRIOR APPLICATION DATA:

APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

## INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-7

Query Match 92.4%; Score 158; DB 3; Length 31;  
Best Local Similarity 90.3%; Pred. No. 5.2e-17;  
Matches 28; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 SRTHRSMEIRTPDINPAMYASRGIRPVGRF 31  
DB 1 SRAHSHSMEIRTPDINPAMYASRGIRPVGRF 31

Search completed: September 13, 2002, 09:20:57  
Job time: 622 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:34 ; Search time 399.68 Seconds  
(without alignments)  
1.945 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_16\_22

Perfect score: 30

Sequence: 1 PVGRFX 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues 747574

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 08

Maximum Match 1008

Listing first 45 summaries

Database :  
1: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDS1/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	93.3	8	22	AA846955
2	28	93.3	10	20	AAW95178
3	28	93.3	15	18	AAW31400
4	28	93.3	15	20	AAW97230
5	28	93.3	15	21	AA149293
6	28	93.3	15	21	AA149295
7	28	93.3	15	21	AAV49296
8	28	93.3	20	18	AAW31394
9	28	93.3	20	18	AAW31387
10	28	93.3	20	18	AAW31374
11	28	93.3	20	20	AAW97232

12	28	93.3	20	20	AAW97234
13	28	93.3	20	20	AAW97236
14	28	93.3	20	20	AAW95191
15	28	93.3	20	20	AAW95175
16	28	93.3	20	21	AAW10350
17	28	93.3	20	21	AAW10358
18	28	93.3	20	21	AAW10365
19	28	93.3	20	21	AAW10369
20	28	93.3	20	21	AAV49294
21	28	93.3	20	21	AAV49301
22	28	93.3	20	21	AAV49302
23	28	93.3	20	22	AAW62519
24	28	93.3	20	22	AAW62527
25	28	93.3	20	22	AAW62534
26	28	93.3	20	22	AAW62538
27	28	93.3	20	22	AAW90992
28	28	93.3	20	22	AAW90994
29	28	93.3	20	22	AAW90996
30	28	93.3	20	22	AAW90997
31	28	93.3	20	22	AAW90998
32	28	93.3	21	18	AAW31395
33	28	93.3	21	18	AAW31388
34	28	93.3	21	18	AAW31375
35	28	93.3	21	20	AAW97227
36	28	93.3	21	20	AAW87616
37	28	93.3	21	20	AAW95192
38	28	93.3	21	21	AAW10351
39	28	93.3	21	21	AAW10359
40	28	93.3	21	21	AAW10366
41	28	93.3	21	22	AAW62520
42	28	93.3	21	22	AAW62528
43	28	93.3	21	22	AAW62535
44	28	93.3	22	18	AAW31396
45	28	93.3	22	18	AAW31389

## ALIGNMENTS

RESULT 1	AAW97234
ID	AAW97234 standard: Protein: 8 AA.
AC	AAW97234
XX	
04-MAY-2001 (first entry)	
DT	
XX	
DE	Peptide PRP8 fragment.
XX	
KW	GPRI0; UHR-1; PRP receptor; prolactin-releasing peptide; pain;
KW	central nervous system disorder; autonomic regulation; analgesic;
KW	hypotensive; blood pressure.
XX	
OS	Unidentified.
XX	
FT	Key
FT	Modified-site
XX	Location/Qualifiers
XX	8
XX	/note="C-terminal amide"
PD	WO200109182-A1.
PD	08-FEB-2001.
XX	
PD	03-AUG-2000; 2000WO-F10064.
XX	
PR	03-AUG-1999; 99US-0365756.
XX	20-MAR-2000; 2000US-0531567.
PA	(JUVVA-) JUVANTIA PHARMA LTD OY.
XX	
PI	Panula PAJ, Petrovarara A, Kalso E, Korpi E;
XX	
DR	WPI: 2001-182941/18.

Rat type ligand po  
Human type ligand  
Bovine pituitary-d  
Murine pituitary-d  
Bovine pituitary-d  
Bovine oxytocin se  
Rat oxytocin secre  
Human oxytocin sec  
Oxytocin secretion  
19P2 ligand peptid  
19P2 ligand peptid  
19P2 ligand peptid  
19P2 ligand peptid  
19P2 ligand peptid  
Bovine CRH releas  
Rat CRH releas  
Human CRH releas  
CRH releasing prot  
Prolactin releas  
Prolactin releas  
Prolactin releas  
Peptide PRP20 fra  
Synthetic G prote  
Human type G prote  
Rat type G prote  
Bovine G protein-c  
Partial ligand pol  
Mammalian 19P2 lig  
Bovine pituitary-d  
Bovine oxytocin se  
Rat oxytocin secre  
Human oxytocin sec  
Bovine CRH releas  
Rat CRH releas  
Human CRH releas  
Human type G prote  
Rat type G prote

XX C-terminal fragments of prolactin-releasing peptide useful for  
PT regulating autonomic functions and in the manufacture of a medicament  
PT for regulating blood pressure -  
XX  
PS Claim 2; page 10; 40pp; English.  
XX  
CC This invention describes a novel C-terminal fragment (I) of an isolated  
CC prolactin-releasing peptide (PrRP), referred to as PrRP20 and having a  
CC sequence (S1). The invention also describes (1) a therapeutic composition  
CC (C1) comprising (I) or a C-terminal fragment of PrRP referred to as PrRP8  
CC and comprising a sequence (S2); (2) a diagnostic method based on antisera  
CC against PrRP20 for identification of disorders involving the central  
CC nervous system, including those associated with pain or autonomic  
CC regulation, where specific antisera against the N-and/or C-terminal  
CC domains of PrRP is used to identify alterations in PrRP synthesis or  
CC levels; (3) a rat or human receptor encoded by a 1122 nucleotide sequence  
CC (S3), fully defined in the specification; (4) treating a person suffering  
CC from a disorder regulated by a receptor (II) encoded by a sequence of  
CC 1122 nucleotide sequence, fully defined in the specification, located in  
CC the central nervous system, by administering an agonist or antagonist to  
CC the receptor; and (5) treating blood pressure, by blocking of receptors  
CC of PrRP or its C-terminal fragment GlyleargProValGlyArgphe-NH<sub>2</sub> (S2).  
CC The products of the invention have analgesic and hypotensive activity.  
CC (I) is useful for regulating autonomic functions, such as increasing  
CC blood pressure. (I) is useful for treating pain, for manufacturing a  
CC medicament for regulating blood pressure, and for treating pain. Agonist  
CC and antagonist of (II) are useful for treating acute pain, inflammatory  
CC pain and neuropathic pain, for regulating autonomic functions and  
CC treating high blood pressure.  
XX  
SO Sequence 8 AA:  
  
Query Match 93.3%; Score 28; DB 22; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
| | | | |  
DB 4 pvgf 8

RESULT 2  
ID AAW95178 standard; Protein: 10 AA.  
AC AAW95178:  
XX  
DT 10-MAR-1999 (first entry)  
XX  
DE Murine pituitary-derived ligand polypeptide antigenic epitope.  
XX  
XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KW gene therapy; transgenic animal; epitope.  
XX  
OS Mus sp.  
XX  
XX WO9849295-A1.  
PN  
XX  
PD 05-NOV-1998.  
XX  
PF 27-APR-1998; 98WO-JP01923.  
XX  
PR 28-APR-1997; 97JP-0109974.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fukusumi S, Hinuma S;  
XX  
XX WPI; 1999-009423/01.  
XX  
PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system; pituitary and  
PT pancreas, and for drug screening  
XX  
PS Disclosure; Page 26; 206pp; English.  
XX  
CC The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.  
XX  
SO Sequence 10 AA:  
  
Query Match 93.3%; Score 28; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 4.5;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
| | | | |  
DB 6 pvgf 10

RESULT 3  
ID AAW31400 standard; Peptide: 15 AA.  
AC AAW31400:  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Synthetic ligand 19P2-L31 peptide II.  
XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent; antigen.  
XX  
OS Synthetic.  
XX  
XX WO9724436-A2.  
PN  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
XX  
PR 28-DEC-1995; 95JP-034371.  
XX  
PR 15-MAR-1996; 96JP-0059419.  
XX  
PR 12-AUG-1996; 96JP-0211805.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
 PI Kawamata Y, Kikada C;  
 XX  
 DR WPI: 1997-363672/33.  
 XX  
 PS Example 43; Page 151; 258pp; English.  
 XX  
 CC This peptide contains the partial C-terminal sequence of the synthetic  
 CC ligand polypeptide 19P2-L31 which is capable of binding to a G  
 CC protein-coupled receptor protein. This peptide is used as an antigen to  
 CC prepare rabbit anti-bovine 19P2-L31 antibodies which are used in binding  
 CC assays. Pharmaceutical compositions containing this ligand may be used  
 CC as a pituitary function modulator, a central nervous system modulator  
 CC or a pancreatic function modulator. This ligand could have specific  
 CC applications as a prophylactic or therapeutic agent for dementia,  
 CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
 CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
 CC hyper- and polydipsia, hyperlipidaemia, hypercholesterolaemia,  
 CC hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
 CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
 CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
 CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
 CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
 CC and/or oligosaccharide. Assays can also be developed to screen compounds  
 CC which are capable of altering the binding activity of the ligand  
 CC affecting activation of the G protein-coupled receptor protein.  
 CC  
 SQ Sequence 15 AA:  
 XX  
 OY 1 PVGRF 5  
 Y 11111  
 Db 11 pvrft 15  
 XX  
 RESULT 4  
 AAM97230 ID AAM97230 standard; Peptide: 15 AA.  
 XX  
 AC AAM97230;  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE C-terminal ligand polypeptide derived antigen.  
 XX  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;  
 KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriorcarcinoma; hydaltid mole;  
 KW interruption mole; abortion; unfertilty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia; prolactin secretion.  
 XX  
 OS Synthetic.  
 XX  
 PN W09858962-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 22-JUN-1998; 98WO-JP02765.  
 XX  
 PR 23-JUN-1997; 97JP-0165437.  
 XX

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 XX  
 DR WPI: 1999-105614/09.  
 XX  
 PS Example 43; Page 123; 241pp; English.  
 XX  
 CC AAM97229-31 represent a ligand polypeptide derived fragments used to  
 CC produce antibodies. The specification describes an agent for modulating  
 CC protein secretion which comprises a ligand polypeptide or a salt, for  
 CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
 CC prolactin secretion can be used for treating or preventing  
 CC hypovarianism, gonocyst cacogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,  
 CC brain tumour, emmenorrhoea, autoimmune disease, prolactinoma,  
 CC infertility, impotence, amenorrhoea, galactorrhea, acromegaly,  
 CC Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright  
 CC syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory  
 CC agents can also be used as contraceptives. The agents for modulating  
 CC placental function can be used for treating or preventing  
 CC choriorcarcinoma, hydaltid mole, abortion, unfertilty  
 CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
 CC oxytocia.  
 CC  
 SQ Sequence 15 AA:  
 XX  
 OY 1 PVGRF 5  
 Y 11111  
 Db 11 pvrft 15  
 XX  
 RESULT 5  
 AAY49293 ID AAY49293 standard; peptide: 15 AA.  
 XX  
 AC AAY49293;  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE 19P2 ligand peptide fragment.  
 XX  
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15  
 FT /note="C-terminal amide"  
 XX  
 PN W09960112-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 20-MAY-1999; 99WO-JP02650.  
 XX  
 PR 21-MAY-1998; 98JP-0140293.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX

PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI: 2000-039381/03.  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality  
XX  
PS Disclosure: Page 26; 73pp; Japanese.  
XX  
CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA;

Query Match 93.3%; Score 28; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 11 pygrf 15

RESULT 6  
AAY49295 standard; peptide; 15 AA.

AC AAY49295;

DT 22-FEB-2000 (first entry)

DE 19p2 ligand peptide fragment.

KM Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
KM pituitary; regulatory mechanism; central nervous system; pancreatic.

OS Homo sapiens.

Key Location/Qualifiers

FT Modified-site 1 /note="N-terminal acetylation"

FT Modified-site 4 /note="acetylated Tyr"

FT Modified-site 15 /note="C-terminal amide"

PN WO960112-A1.

PT 25-NOV-1999.

PF 20-MAY-1999; 99WO-JP02650.

PR 21-MAY-1998; 98JP-0140293.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

XX WPI: 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
XX studying diseases related to ligand abnormality  
XX  
XX Disclosure: Page 26; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA;

Query Match 93.3%; Score 28; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 11 pygrf 15

RESULT 7  
AAY49296 standard; peptide; 15 AA.

AC AAY49296;

DT 22-FEB-2000 (first entry)

DE 19p2 ligand peptide fragment.

KM Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
KM pituitary; regulatory mechanism; central nervous system; pancreatic.

OS Homo sapiens.

Key Location/Qualifiers

FT Modified-site 15 /note="C-terminal amide"

PN WO960112-A1.

PT 25-NOV-1999.

PF 20-MAY-1999; 99WO-JP02650.

PR 21-MAY-1998; 98JP-0140293.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

XX WPI: 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
XX studying diseases related to ligand abnormality  
XX  
XX Disclosure: Page 27; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA;



Query Match 93.3%; Score 28; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 6.7;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
Db 11 pvgfr 15

RESULT 8

AAW31394  
ID AAW31394 standard; Peptide; 20 AA.  
XX  
AC AAW31394;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96GP-0246573.  
PR 28-DEC-1995; 95GP-0343371.  
PR 15-MAR-1996; 96GP-0059419.  
PR 12-AUG-1996; 96GP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kltada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02431.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polypnea,  
CC hyperlipidemia, hypercholesterolemia, hyperglycemia,  
CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 93.3%; Score 28; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
Db 16 pvgfr 20

RESULT 9

AAW31387  
ID AAW31387 standard; Peptide; 20 AA.  
XX  
AC AAW31387;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Rat type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Rat sp.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96GP-0246573.  
PR 28-DEC-1995; 95GP-0343371.  
PR 15-MAR-1996; 96GP-0059419.  
PR 12-AUG-1996; 96GP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kltada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02424.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 180; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the  
CC sequence represented in AAW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polypnea,  
CC hyperlipidemia, hypercholesterolemia, hyperglycemia,  
CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 93.3%; Score 28; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.9;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
Db 16 pvgf 20

RESULT 10  
AAW31374  
ID AAW31374 standard; Peptide: 20 AA.

XX AAW31374;

XX 06-APR-1998 (first entry)

XX Bovine G protein-coupled receptor ligand peptide fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical;

XX modulator; pituitary; central nervous system; pancreas; prolactin;

XX therapeutic agent.

XX Bos taurus.

XX MO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96MO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

XX 28-DEC-1995; 95JP-0343371.

XX 15-MAR-1996; 96JP-0059419.

XX 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

XX Kawamata Y, Kikada C;

XX WPI: 1997-363672/33.

XX N-PSDB: AAV02397.

XX Claim 2; Page 161; 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary  
XX derived ligand corresponding to amino acid residues 34 to 53 of the  
XX sequence in AAW31368 and is used in an assay to monitor ligand binding  
XX to the G protein-coupled receptor protein. Pharmaceutical compositions  
XX containing this ligand may be used as a pituitary function modulator, a  
XX central nervous system modulator or a pancreatic function modulator.  
XX This ligand could have specific applications as a prophylactic or  
XX therapeutic agent for dementia, depression, hyperkinetic syndrome,  
XX disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
XX growth hormone secretory disease, hyper- and polypnea,  
XX hypercholesterolemia, hyperglycemia, hyperlipidemia,  
XX hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
XX Turner's syndrome, neurosis, rheumatoid arthritis, spinal injury,  
XX transient brain ischemia, amyotrophic lateral sclerosis, acute  
XX myocardial infarction, spinocerebellar degeneration, bone fracture,  
XX trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility  
XX and/or oligosacchara. Assays can also be developed to screen compounds  
XX which are capable of altering the binding activity of the ligand thus  
XX affecting activation of the G protein-coupled receptor protein.

XX Sequence 20 AA;

Query Match 93.3%; Score 28; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
Db 16 pvgf 20

RESULT 11  
AAW97232  
ID AAW97232 standard; peptide: 20 AA.

XX AAW97232;

XX 06-MAY-1999 (first entry)

XX Bovine pituitary-derived ligand polypeptide fragment.

XX Bovine pituitary-derived ligand; modulation; prolactin secretion;

XX G protein-coupled receptor; GPCR; hypovarianism; gonocyst cecogenesis;

XX menopausal syndrome; euthyroid; hypometabolism; lactation;

XX pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;

XX prolactinoma; infertility; impotence; amenorrhea; galactorrhea;

XX acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;

XX Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;

XX contraceptive; placental function; chorioncarcinoma; hydatid mole;

XX abortion; unthrifty fetus; abnormal saccharometabolism;

XX abnormal lipidmetabolism; oxytocia.

XX Bos sp.

XX MO9858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98MO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI: 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin

XX secretion or placental function, e.g. for treating menopausal

XX syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3; Page 136; 241pp; English.

XX The present sequence represents a bovine pituitary-derived ligand  
XX fragment. It is used in the course of the invention. The specification  
XX describes an agent for modulating prolactin secretion which comprises a  
XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
XX protein. The agents for promoting prolactin secretion can be used for  
XX treating or preventing hypovarianism, gonocyst cecogenesis, menopausal  
XX syndrome, euthyroid or hypometabolism. They can be used for promoting  
XX lactation in a domestic mammal and as an aphrodisiac. The agents for  
XX inhibiting prolactin secretion can be used for treating or preventing  
XX pituitary adenomatosis, brain tumour, amenorrhea, galactorrhea,  
XX prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
XX acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
XX Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dysospermia.  
XX The inhibitory agents can also be used as contraceptives. The agents for  
XX modulating placental function can be used for treating or preventing  
XX chorioncarcinoma, hydatid mole, abortion, unthrifty fetus,  
XX abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 20 AA;

Query Match 93.3%; Score 28; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
Db 16 pvgrf 20

## RESULT 12

AAW97234  
ID AAW97234 standard; peptide: 20 AA.  
XX  
AC AAW97234;  
XX

DT 06-MAY-1999 (first entry)  
XX

DE Rat type ligand polypeptide fragment.  
XX

KM Rat type ligand: modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacogenesis;  
KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KM contracteptive; placental function; choriocarcinoma; hydatid mole;  
KM interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KM abnormal lipidmetabolism; oxytocia.  
XX

OS Rattus sp.  
XX

PN MO9858962-A1.  
XX

PD 30-DEC-1998.  
XX

PF 22-JUN-1998; 98MO-JP02765.  
XX

PR 23-JUN-1997; 97JP-0165437.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX

DR WPI; 1999-105614/09.  
XX

PT Use of G protein-coupled receptor ligands - for modulating prolactin  
secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 154; 241pp; English.  
XX

CC The present sequence represents a rat type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonecyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX

SQ Sequence 20 AA;  
XX

Query Match 93.3%; Score 28; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
Db 16 pvgrf 20

## RESULT 13

AAW97236  
ID AAW97236 standard; peptide: 20 AA.  
XX  
AC AAW97236;  
XX

DT 06-MAY-1999 (first entry)  
XX

DE Human type ligand polypeptide fragment.  
XX

KM Rat type ligand: modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacogenesis;  
KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KM contracteptive; placental function; choriocarcinoma; hydatid mole;  
KM interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KM abnormal lipidmetabolism; oxytocia.  
XX

OS Homo sapiens.  
XX

PN MO9858962-A1.  
XX

PD 30-DEC-1998.  
XX

PF 22-JUN-1998; 98MO-JP02765.  
XX

PR 23-JUN-1997; 97JP-0165437.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX

DR WPI; 1999-105614/09.  
XX

PT Use of G protein-coupled receptor ligands - for modulating prolactin  
secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 166; 241pp; English.  
XX

CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonecyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX

SQ Sequence 20 AA;  
XX

Query Match 93.3%; Score 28; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYGRF 5  
11111  
Db 16 pygrf 20

## RESULT 14

AAM95191  
ID AAM95191 standard: peptide; 20 AA.

AC AAM95191;

DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; bovine.

XX Bos sp.

XX W09849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hinuma S;

XX WPI: 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

XX Example 19; Page 151; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of the  
CC polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. The present sequence represents a bovine genome-derived ligand  
CC polypeptide fragment which is similar to the murine ligand-polypeptide.  
XX Sequence 20 AA;

Query Match 93.3%; Score 28; DB 20; Length 20;

Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 PYGRF 5  
11111  
Db 16 pygrf 20

## RESULT 15

AAM95175  
ID AAM95175 standard: Protein; 20 AA.

AC AAM95175;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; epitope.

XX Mus sp.

XX W09849295-A1.

XX 05-NOV-1998.

XX 27-APR-1998; 98WO-JP01923.

XX 28-APR-1997; 97JP-0109974.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fukusumi S, Hinuma S;

XX WPI: 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

XX Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. Sequences AAM95174 to AAM95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.  
XX Sequence 20 AA;

SQ Query Match 93.3%; Score 28; DB 20; Length 20;

Query Match 93.3%; Score 28; DB 20; length 20;  
Best Local Similarity 100.0%; Pred. No. 8.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 16 PVGRF 20

Search completed: September 13, 2002, 09:18:34  
Job time: 499 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:16 ; Search time 311.85 seconds  
(without alignments)  
12.204 Million cell updates/sec

Title: US-09-446-543a-73  
Perfect score: 109  
Sequence: 1 TPDINPAMYXXRGIRPVGRFX 22

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_prodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	75.2	117	13	09W624 carassius a
2	49.5	45.4	664	16	0911W4 pseudomonas
3	49	45.0	790	10	09M371 arabidopsis
4	46.5	42.7	420	2	0931E7 streptomyces
5	46	42.2	54	4	090U99 xylella fas
6	46	42.2	333	16	09PH76 homo sapien
7	46	42.2	465	4	060687 homo sapien
8	45	42.2	540	10	091G20 oryza sativ
9	45	41.3	767	5	020170 caenorhabd
10	44	40.4	145	5	095ZB5 leishmania
11	44	40.4	250	16	0912F2 pseudomonas
12	44	40.4	284	17	050128 pyrococcus
13	44	40.4	419	4	09Y276 homo sapien
14	43.5	39.9	1501	10	09SD86 arabidopsis
15	43	39.4	220	16	091022 pseudomonas
16	43	39.4	232	16	09R221 delnoccocus

17	43	39.4	309	5	017234 caenorhabd
18	43	39.4	333	4	096SD4 homo sapien
19	43	39.4	370	11	09D3V7 mus musculu
20	43	39.4	418	11	09CXP5 mus musculu
21	43	39.4	1488	5	020294 caenorhabd
22	42.5	39.0	333	2	09RJ10 streptomyces
23	42.5	39.0	360	5	019879 caenorhabd
24	42	38.5	293	2	053855 spiroplasma
25	42	38.5	416	16	099ZAG9 streptococ
26	42	38.5	428	16	09PKET chlamydia m
27	42	38.5	433	16	09A382 caulobacter
28	42	38.5	349	2	09EW96 streptomyces
29	41.5	38.1	151	12	091EQ1 dashen mos
30	41.5	38.1	345	16	09XGV2 bacillus ha
31	41	37.6	97	2	033440 pseudomonas
32	41	37.6	128	16	P74747 synchocyst
33	41	37.6	226	16	09A359 caulobacter
34	41	37.6	280	4	096925 homo sapien
35	41	37.6	284	17	09VQ03 pyrococcus
36	41	37.6	328	4	09NVR5 homo sapien
37	41	37.6	338	5	018729 caenorhabd
38	41	37.6	343	2	09ZA30 streptomyces
39	41	37.6	414	2	033480 propionibac
40	41	37.6	545	16	09A/W7 caulobacter
41	41	37.6	547	2	055020 streptomyces
42	41	37.6	547	2	093H76 streptomyces
43	41	37.6	2155	11	008523 mus musculu
44	40.5	37.2	316	11	09EPF7 mus musculu
45	40.5	37.2	779	3	09P5J9 neurospora

#### ALIGNMENTS

RESULT	ID	Query Match	Best Local Similarity	Matches	Score	Pred. No.	Length	Indels	Gaps
1	09W624	75.2%	60.0%	12	82	4e-06	117	0	0
09W624	PREDIMINARY;								
AC	09W624;								
DT	01-NOV-1999 (TREMUREL. 12, Created)								
DT	01-NOV-1999 (TREMUREL. 12, Last sequence update)								
DT	01-DEC-2001 (TREMUREL. 19, Last annotation update)								
DE	C-RE AMIDE.								
OS	Carassius auratus (Goldfish).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;								
OC	Cypriniformes; Cyprinidae; Carassius.								
OX	NCBI_TaxID=7957;								
RN	[1]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=BRAIN;								
RA	Satoh H., Minakata H., Fujimoto M.;								
RT	"Carassius auratus (C-RE amide)";								
RL	Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.								
DR	EMBL; AB020024; BAA7662.1;								
SQ	SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;								
Query Match		75.2%							
Best Local Similarity		60.0%							
Matches	12; Conservative	5;	Mismatches	3;	Indels	0;	Gaps	0;	
QY	1 TPDINPAMYXXRGIRPVGRF 20								
DB	56 SPEDIPFMYVGRVPIGRF 75								
RESULT	2								
ID	0911W4								
AC	0911W4								
DT	01-MAR-2001 (TREMUREL. 16, Created)								
DT	01-MAR-2001 (TREMUREL. 16, Last sequence update)								
DT	01-OCT-2001 (TREMUREL. 18, Last annotation update)								

DE HYPOTHETICAL PROTEIN PA2151.  
GN PA2151.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AVCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
RA Garber R.L., Goltz L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT \*Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen\*;  
RL Nature 406:959-964(2000).  
DR EMBL: AE004642; AGO5539.1; -  
DR InterPro: IPR001589; Actinln\_act\_bind.  
DR InterPro: IPR000461; Alpha\_amyase.  
DR Pfam: PF00128; alpha-amyase; 1.  
DR PROSITE: PS00019; ACTININ.1; UNKNOWN.1.  
KM Hypothetical protein: Complete proteome.  
SQ SEQUENCE 664 AA; 76329 MW; 8F59FED54C308AD CRC64;

Query Match 45.4%; Score 49.5; DB 16; Length 664;  
Best Local Similarity 62.5%; Pred. No. 6.5;  
Matches 10; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 TPDINPAWYXXRGIRP 16  
DB 478 TPDINP-WFLQSGRP 492

RESULT 3  
ID 09M371 PRELIMINARY; PRT; 790 AA.  
AC 09M371;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 87.4 KDA PROTEIN.  
GN F15616.60.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Queller F., Salanoubat M.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL13959; CAB71097.1; -  
KM Hypothetical protein.  
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 45.0%; Score 49; DB 10; Length 790;  
Best Local Similarity 47.4%; Pred. No. 9.5;  
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 PDINPAWYXXRGIRPVGRF 20  
DB 366 PPHNPRYSGRGQPHGRW 384

RESULT 4  
ID 093L27 PRELIMINARY; PRT; 420 AA.  
AC 093L27;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CCM3239.  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT \*Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CCM3239.\*;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY033994; AAK61719.1; -  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22B88C2DEA CRC64;

Query Match 42.7%; Score 46.5; DB 2; Length 420;  
Best Local Similarity 52.6%; Pred. No. 13;  
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

OY 7 AWYXX-----RGIRPVGRF 20  
DB 37 AWMAAVLRGSGSGIRPVGRF 55

RESULT 5  
ID 09UJF9 PRELIMINARY; PRT; 54 AA.  
AC 09UJF9;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
DE D1479J7.3 (SUSHI-REPEAT PROTEIN (SRPU)) (FRAGMENT).  
GN D1479J7.3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Lawlor S.;  
RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL035608; CAB55682.1; -  
FT NON\_TER 54  
SQ SEQUENCE 54 AA; 6110 MW; E2F3C39F7B961A9F CRC64;

Query Match 42.2%; Score 46; DB 4; Length 54;  
Best Local Similarity 43.8%; Pred. No. 1.8;  
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 TPDINPAWYXXRGIRP 16  
DB 18 TPAVPTWYAGSGYYP 33

RESULT 6  
ID 09PH76 PRELIMINARY; PRT; 333 AA.  
AC 09PH76;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYDROXYBENZONATE OCTAPRENTITRANSFERASE.



GN XF0068.  
OS *Xylella fastidiosa*.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC *Xylella*.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britton M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Honkeisel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Niani A.J., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA de Oliveira M.C., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tsunako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Melandri J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen *Xylella fastidiosa*.";  
RL Nature 406:151-159(2000).  
DR EMBL: AE003860; AAF82881.1; -.  
DR InterPro: IPR000537; UblA.  
DR Pfam: PF01040; UblA; 1.  
DR PROSITE: PS00943; UblA; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 333 AA; 37931 MW; ECF3P4716C962B95 CRC64;

Query Match 42.2%; Score 46; DB 16; Length 333;  
Best Local Similarity 53.3%; Pred. No. 12;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 4 INPAMYXXRGIRPVG 18  
Db 54 LDPMYKRLARDRPVG 68

RESULT 7  
O60687  
ID 060687 PRELIMINARY; PRT; 465 AA.  
AC 060687;  
DT 01-AUG-1998 (TrEMBLrel. 07, Created)  
DT 01-AUG-1998 (TrEMBLrel. 07, last sequence update)  
DE 01-JUN-2001 (TrEMBLrel. 17, last annotation update)  
DE SUSH1-REPEAT PROTEIN.  
DE SRPUL.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Kurosawa H., Inukai T., Inaba T., Gol K., Chang K.-S., Sinjo T.,  
RA Rakestraw K.M., Naeve C.W., Look T.A.;  
RL Submitted (APR-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF060567; AAC15765.1; -.

DR InterPro: IPR001128; Cyt\_P450.  
DR InterPro: IPR003410; HYR.  
DR InterPro: IPR000436; Sush1\_SCR\_CCP.  
DR Pfam: PF02494; HYR; 1.  
DR Pfam: PF00084; sush1; 3.  
DR SMART: SM00032; CCP; 3.  
DR PROSITE: PS00086; CYTOCHROME\_P450; UNKNOWN\_1.  
SQ SEQUENCE 465 AA; 52971 MW; 4D752B187FE3EFB8 CRC64;

Query Match 42.2%; Score 46; DB 4; Length 465;  
Best Local Similarity 43.8%; Pred. No. 17;  
Matches 7; Conservative 1; Mismatches 8; Indels 0; Gaps 0;

OY 1 TPQINPAMYXXRGIRP 16  
Db 18 TPATPTWYAGSGYRP 33

RESULT 8  
O9LGZ0  
ID 09LGZ0 PRELIMINARY; PRT; 540 AA.  
AC 09LGZ0;  
DT 01-OCT-2000 (TrEMBLrel. 15, Created)  
DT 01-OCT-2000 (TrEMBLrel. 15, last sequence update)  
DE 01-DEC-2001 (TrEMBLrel. 19, last annotation update)  
DE EST AV070209(R3722) CORRESPONDS TO A REGION OF THE PREDICTED  
DE GENE.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Eriarthroideae; Oryzeae; Oryza.  
OX NCBI\_TaxID=4530;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CV. NIPPONBARE;  
RA Sasaki T., Matsumoto T., Yamamoto K.;  
RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC  
RT clone:p0702f03.";  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER  
CC PROTEINS (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +  
CC PYROPHOSPHATE + PROTEIN N-UBIQUITINYLLYSINE.  
CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.  
CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-  
CC THIOLESTER FORMATION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
DR EMBL: AP002481; BAA96583.1; -.  
DR HSSP: P06104; IAYZ.  
DR InterPro: IPR000608; UBQ\_conjugat.  
DR Pfam: PF00179; UBQ\_con; 1.  
DR SMART: SM00212; UBQC; 1.  
DR PROSITE: PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
KW Ligase; Ubiquitin conjugation.  
SQ SEQUENCE 540 AA; 60487 MW; 5DE1FAEEB75A86E CRC64;

Query Match 42.2%; Score 46; DB 10; Length 540;  
Best Local Similarity 50.0%; Pred. No. 20;  
Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 3 DINPAMYXXRGIRP 16  
Db 69 DLGVAMRWVGLRP 82

RESULT 9  
O20170  
ID 020170 PRELIMINARY; PRT; 767 AA.  
AC 020170;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, last sequence update)

DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
DE F38E11.7 PROTEIN.  
GN F38E11.7.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Chromadorea; Rhabditida; Rhabditoidea;  
OC Rhabditidae; Pelodermidae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Mathews P.;  
RL Submitted (JAN-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology.";  
RL Science 282:2012-2018(1998).  
DR EMBL: Z68342; CAA92775.1; -.  
DR InterPro: IPR000636; Cation\_chan\_non\_119.  
DR InterPro: IPR001622; Channel\_pore\_K.  
DR InterPro: IPR000595; CNMP\_binding.  
DR Pfam: PF00027; CNMP\_binding.1.  
DR Pfam: PF00520; ion\_trans.1.  
DR SMART: SM00100; CNMP.1.  
DR PROSITE: PS00888; CNMP\_BINDING\_1; UNKNOWN\_1.  
DR PROSITE: PS50042; CNMP\_BINDING\_3; 1.  
DR PROSITE: PS50042; CNMP\_BINDING\_3; 1.  
SQ SEQUENCE 767 AA; 89988 MW; F7ECF69DBEACFC3 CRC64;

Query Match 41.3%; Score 45; DB 5; Length 767;  
Best Local Similarity 75.0%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 PDINPAMY 9  
II:||||  
DB 747 PDVCPAMY 754

RESULT 10  
ID 095ZB5 PRELIMINARY; PRT; 145 AA.  
AC 095ZB5;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
DE 01-DEC-2001 (TREMblrel. 19, last annotation update)  
DE HYPOTHETICAL 15.8 KDA PROTEIN.  
GN L1994.09.  
OS Leishmania major.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_TaxID=5664;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FRIDLIN;  
RA Schaefer M., Mueller-Auer S., Gabel C., Fuchs M., Rieger M.,  
Ivens A.C., Quail M., Rajandream M.A., Bartell B.G.;  
RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=FRIDLIN;  
RX MEDLINE=98146435; PubMed=9477341;  
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,  
Smith D.F.;  
RT "A physical map of the Leishmania major Friedlin genome.";  
RL Genome Res. 8:135-145(1998).  
DR EMBL: AL596272; CAC44727.1; -.  
KW Hypothetical protein.  
SQ SEQUENCE 145 AA; 15806 MW; F30454BC299BAF83 CRC64;

Query Match 40.4%; Score 44; DB 5; Length 145;  
Best Local Similarity 40.0%; Pred. No. 11;  
Matches 8; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

OY 1 TPDINPAMYXXRCIRPVGR 20  
II:|||| I: I I:  
DB 6 TPEANPALALRKVHPAAHW 25

RESULT 11  
ID 0912F2 PRELIMINARY; PRT; 250 AA.  
AC 0912F2;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, last annotation update)  
DE HYPOTHETICAL PROTEIN PA1952.  
PA1952.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stoyer C.K., Pham X.O.T., Ertwin A.L., Mizoguchi S.D., Warren P.,  
Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
Gardner R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Llm R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL: AE004622; AAG05340.1; -.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 250 AA; 25619 MW; B997F6BE28D792C2 CRC64.

Query Match 40.4%; Score 44; DB 16; Length 250;  
Best Local Similarity 47.1%; Pred. No. 19;  
Matches 8; Conservative 2; Mismatches 7; Indels 0; Gaps 0;

OY 2 PDINPAMYXXRCIRPVG 18  
I:|||| I: I I:  
DB 233 PALNCAMEOLRLRPSG 249

RESULT 12  
ID 050128 PRELIMINARY; PRT; 284 AA.  
AC 050128;  
DT 01-JUN-1998 (TREMblrel. 06, Created)  
DT 01-JUN-1998 (TREMblrel. 06, last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, last annotation update)  
DE HYPOTHETICAL 32.3 KDA PROTEIN PH1420.  
GN PH1420.  
OS Pyrococcus horikoshii.  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.  
OX NCBI\_TaxID=53953;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=OT3;  
RX MEDLINE=98344137; PubMed=9679194;  
RA Kawarabayashi Y., Sawada M., Horikawa H., Haikawa Y., Hino Y.,  
Yamamoto S., Sekine M., Baba S.-I., Kosugi H., Hosoyama A., Nagai Y.,  
RA Sakai M., Ogura K., Otsuka R., Nakazawa H., Takamiya M., Ohfuku Y.,  
RA Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kushiida N., Oguchi A.,  
RA Aoki K.-I., Yoshizawa T., Nakamura Y., Kikuchi H.;  
RT "Complete sequence and gene organization of the genome of a hyper-  
thermophilic archaeobacterium, Pyrococcus horikoshii OT3.";  
RL DNA Res. 5:55-76(1998).  
DR EMBL: AP000006; BAA30526.1; -.  
DR InterPro: IPR001330; Prenyltrans.

DR InterPro: IPR003662; sub\_transporter.  
DR Pfam: PF00432; prenyltans; 2.  
DR PROSITE: PS00216; SUGAR\_TRANSPORT\_1; UNKNOWN.1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 284 AA; 32319 MW; BE0E7BC3711D3B15 CRC64;

Query Match 40.4%; Score 44; DB 17; Length 284;  
Best Local Similarity 44.4%; Pred. No. 22;  
Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;  
QY 2 PDINPAMYXXKGIKRVGR 19  
Db 217 PYIEPTFYALRGLELGR 234

RESULT 13  
QY276 PRELIMINARY; PRT; 419 AA.  
AC Q9Y276;  
DT 01-NOV-1999 (TREMBlrel. 12, Created)  
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE H-BCS1 (BCS1 (YEAST HOMOLOG)-LIKE).  
GN BCS1 OR BCS1L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=99097350; PubMed=9878253;  
RA Petruzzella V., Tiranelli V., Fernandez P., Ianna P., Carrozzo R.,  
RA Zeviani M.;  
RT "Identification and characterization of human cDNAs specific to BCS1,  
RT PPT112, SCO1, COX15, and COX11, five genes involved in the formation  
RT and function of the mitochondrial respiratory chain.";  
RL Genomics 54:494-504(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=96207227; PubMed=8619474;  
RA Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;  
RT "A 'double adaptor' method for improved shotgun library  
RT construction.";  
RL Anal. Biochem. 236:107-113(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE=97264341; PubMed=9110174;  
RA Yu W., Andersson B., Morley K.C., Muzny D.M., Ding Y., Liu W.,  
RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;  
RT "Large-scale concatenation cDNA sequencing.";  
RL Genome Res. 7:353-358(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA de Lonlay P., Valnot I., Barrientos A., Gorbatyuk M., Tzagoloff A.,  
RA Benayoun E., Chretien D., Kadhom N., Lombes A., Ogier de Baulny H.,  
RA Maudet P., Munnich A., Rustin P., Rotig A.;  
RT "Mutations in bcs1, a mitochondrial respiratory chain assembly gene,  
RT are responsible for the complex III deficiency of patients with  
RT tubulopathy and liver failure.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE-RHABDOMYOSARCOMA;  
RA Strausberg R.;  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE-MUSCLE, RHABDOMYOSARCOMA;  
RA Strausberg R.;

RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF026849; AAD08638.1; -  
DR EMBL: AF038195; AAB97365.1; -  
DR EMBL: AF346835; AAK29417.1; -  
DR EMBL: BC000416; AAH00416.1; -  
DR EMBL: BC007500; AAH07500.1; -  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003959; AAA\_subfam.  
DR Pfam: PF00004; AAA.1.  
DR SMART: SM00382; AAA.1.  
KW ATP-binding; Hypothetical protein.  
SQ SEQUENCE 419 AA; 47534 MW; 7F0F98BA62F2CBB8 CRC64;

Query Match 40.4%; Score 44; DB 4; Length 419;  
Best Local Similarity 70.0%; Pred. No. 33;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;  
QY 5 NPAMYXXKGI 14  
Db 211 NPKWYTDKGI 220

RESULT 14  
QYSD86 PRELIMINARY; PRT; 1501 AA.  
AC QYSD86;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAR-2001 (TREMBlrel. 16, Last annotation update)  
DE HYPOTHETICAL 164.4 KDA PROTEIN (GENOMIC DNA, CHROMOSOME 5, P1  
DE CLONE:MXM12).  
GN F13G24.180.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Beyan M., Van Der Schueren J., Chuang Y.J., Voet M., Robben J.,  
RA Volckaert G., Bancroft I., Mewes H.W., Lemcke K., Meyer K.F.X.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-COLUMBIA;  
RX MEDLINE=97471969; PubMed=9330910;  
RA Sato S., Kotani H., Nakamura Y., Kaneko T., Asamizu E., Fukami M.,  
RA Miyajima N., Tabata S.;  
RT "Structural analysis of Arabidopsis thaliana chromosome 5. I. Sequence  
RT features of the 1.6 mb regions covered by twenty physically assigned  
RT pl clones.";  
RL DNA Res. 4:215-230(1997).  
DR EMBL: AL133421; CAB62610.1; -  
DR EMBL: AB005249; BAB09962.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 1501 AA; 164360 MW; EAD2AE0C9E8BE5D2 CRC64;

Query Match 39.9%; Score 43.5; DB 10; Length 1501;  
Best Local Similarity 34.5%; Pred. No. 1.5e+02;  
Matches 10; Conservative 5; Mismatches 5; Indels 9; Gaps 2;  
QY 1 TPDIAPAMYXXRG-----IRPV---GRF 20  
Db 1184 SPQMAPSMYSQYGFKNGLVQPVNDTRF 1212

RESULT 15

091022  
 ID 091022 PRELIMINARY; PRT; 220 AA.  
 AC 091022;  
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
 DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
 DE PROBABLE GLUTATHIONE S-TRANSFERASE.  
 GN PA2821.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_Taxid=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964 (2000).  
 DR EMBL: AE004709; AAC06209.1; -.  
 DR InterPro: IPR004046; GST\_LC.  
 DR InterPro: IPR004045; GST\_N.  
 KW Transferase; Complete proteome.  
 SQ SEQUENCE 220 AA; 24716 MW; 6596183EACCA050 CRC64;

Query Match 39.4%; Score 43; DB 16; Length 220;  
 Best Local Similarity 64.3%; Pred. No. 25;  
 Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;  
 OY 6 PAMYXXRGIRPVGR 19  
 ||||| | | | | |  
 Db 39 PAMY--REISPLGR 50

Search completed: September 13, 2002, 09:29:17  
 Job time: 1062 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:42 ; Search time 80.21 Seconds

(without alignments)  
10.620 Million cell updates/sec

Title: US-09-446-543a-73

Perfect score: 109

Sequence: 1 TPDIHPAMYXXRGIRPVGRFX 22

Scoring table: BLOSUM62

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Database: SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	105	96.3	87	1	PRRP_HUMAN
2	104	95.4	83	1	PRRP_RAT
3	104	95.4	98	1	PRRP_BOVIN
4	46	42.2	676	1	EXLI_HUMAN
5	43	39.4	376	1	OPSI_LIMPO
6	43	39.4	376	1	OPSI_LIMPO
7	43	39.4	383	1	CYCR_CHRVI
8	43	39.4	476	1	YAAI_ECOLI
9	43	39.4	719	1	NRPI_YEAST
10	42	38.5	402	1	EXTL_STRCO
11	41.5	38.1	345	1	ARGC_BACHD
12	41	37.6	342	1	Y762_METJA
13	41	37.6	347	1	Y576_METJA
14	41	37.6	986	1	CYCR_ARBPV
15	40	36.7	329	1	Y493_MCTU
16	40	36.7	546	1	CHOD_STRSQ
17	40	36.7	581	1	POL_MLVAK
18	40	36.7	697	1	LCFF_HUMAN
19	40	36.7	843	1	POL_MLVAK
20	40	36.7	1024	1	POPC_RALSO
21	40	36.7	1196	1	POL_MLVAK
22	40	36.7	1196	1	POL_MLVAK
23	39.5	36.2	860	1	VG12_BPOB3
24	39	35.8	149	1	ENRN_BPT7
25	39	35.8	360	1	MNT2_HUMAN
26	39	35.8	377	1	OPSI_HEKSA
27	39	35.8	377	1	OPSI_HEKSA
28	39	35.8	485	1	SYE_BACHD
29	39	35.8	622	1	PPCC_RAT
30	39	35.8	953	1	SYV_VIRCH
31	39	35.8	962	1	UVRA_MERTH
32	38.5	35.3	240	1	PLSC_HELTP
33	38.5	35.3	265	1	UBIE_RICPR

34	38.5	35.3	831	1	DPO1_THERFL
35	38.5	35.3	877	1	CAD2_BOVIN
36	38.5	35.3	906	1	CAD2_HUMAN
37	38.5	35.3	906	1	CAD2_MOUSE
38	38.5	35.3	906	1	CAD2_RAT
39	38.5	35.3	3083	1	POIG_ZYMRV
40	38	34.9	158	1	RLIS_ARPE
41	38	34.9	197	1	ILIB_HUMAN
42	38	34.9	261	1	PHSC_ECOLI
43	38	34.9	288	1	Y587_PASW
44	38	34.9	332	1	LYTB_MYCLE
45	38	34.9	399	1	HMPA_BACSU

## ALIGNMENTS

RESULT 1  
ID PRRP\_HUMAN STANDARD; PRT; 87 AA.  
AC P81277;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRP31; Proactin-releasing peptide PRP20].  
DE PRH.  
GN Homo sapiens (human).  
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID:9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA MEDLINE=98268781; Pubmed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RT "A proactin-releasing peptide in the brain.";  
RT Nature 393:272-276(1998).  
RN [2]  
RP TISSUE SPECIFICITY.  
RX Pubmed=10498338;  
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RT "Tissue distribution of proactin-releasing peptide (PRP) and its receptor.";  
RT Regul. Pept. 83:1-10(1999).  
CC -1- FUNCTION: Stimulates proactin (PR) release and regulates the expression of proactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PR.  
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
CC -----  
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: AB015419; BAA29027.1; -  
CC DR MIM: 602663; -  
KW Hormone; Amidation; Signal.  
FT SIGNAL 1 22  
FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRP31.  
FT PEPTIDE 34 53 PROACTIN-RELEASING PEPTIDE PRP20.  
FT MOD RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 87 AA: 9659 MW: 229A2F35FCF981B CRC64;



RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
RT "Identification and localization of the gene for EXT1, a third member  
RN of the multiple exostoses gene family.";  
RL Genome Res. 7:10-16(1997).  
[2]  
RP SEQUENCE FROM N.A.  
RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;  
RT "Mutations of the EXT genes in hereditary multiple exostoses in  
RL Chinese.";  
RN Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
[3]  
RP SEQUENCE FROM N.A.  
RA Wuyts W., Spleker N., Van Roy N., De Paepe A., De Bouille K.,  
RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;  
RT "Refined physical mapping and genomic structure of the EXT1 gene";  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
CC reticulum (by similarity).  
CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
-----  
DR EMBL: U67191; AAC51141.1; -;  
DR EMBL: AF083633; AAD02840.1; -;  
DR EMBL: AF083623; AAD02840.1; JOINED.  
DR EMBL: AF083624; AAD02840.1; JOINED.  
DR EMBL: AF083625; AAD02840.1; JOINED.  
DR EMBL: AF083626; AAD02840.1; JOINED.  
DR EMBL: AF083627; AAD02840.1; JOINED.  
DR EMBL: AF083628; AAD02840.1; JOINED.  
DR EMBL: AF083629; AAD02840.1; JOINED.  
DR EMBL: AF083630; AAD02840.1; JOINED.  
DR EMBL: AF083631; AAD02840.1; JOINED.  
DR EMBL: AF083632; AAD02840.1; JOINED.  
DR EMBL: AF153980; AAF73172.1; -;  
DR EMBL: AF153981; AAF73172.1; JOINED.  
DR MIM: 601738; -;  
DR InterPro: IPR004263; Exostosin.  
DR Pfam: PF03016; Exostosin; 1.  
KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.  
FT TRANSMEM 10 30  
FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
FT (POTENTIAL).  
SQ SEQUENCE 676 AA; 74673 MM; B5E006A87625633 CRC64;  
FT CARBOHYD 269 269  
FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
Query Match 42.2%; Score 46; DB 1; Length 676;  
Best Local Similarity 45.0%; Pred. No. 4.8;  
Matches 9; Conservative 3; Mismatches 8; Indels 0; Gaps 0;  
Qy 1 TPDIHPAMYXXRGIRPVGRF 20  
Db 400 SPODEFFYYLQGSRRPGRF 419  
RESULT 5  
OPSL LIMPO STANDARD: PRT: 376 AA.  
AC P35360;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Lateral eye opsin.  
OS Limulus polyphemus (Atlantic horseshoe crab).  
CC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
CC Limulidae; Limulus.

OX NCBI\_TaxID=6650;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lateral eye;  
RX MEDLINE=93317641; PubMed=8327495;  
RA Smith W.C., Price D.A., Greenberg R.M., Batteille B.-A.;  
RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
RL Limulus polyphemus";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
CC LINKED TO CIS-RETINAL.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: LATERAL EYE.  
CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
CC BE PHOSPHORYLATED (BY SIMILARITY).  
CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 520 NM.  
CC -1- SIMILARITY: BELONGS TO FAMILY I OF G-PROTEIN COUPLED RECEPTORS.  
CC OPSIN SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).  
-----  
DR EMBL: L03791; AAA28273.1; -;  
DR EMBL: L03781; AAA02498.1; -;  
DR PIR: B48197; B48197.  
DR GCRNB: GCR\_0584; -;  
DR InterPro: IPR000276; GPCR\_Rhodopsn.  
DR Pfam: PF00001; 7tm\_1; 1.  
DR PRINTS: PR00237; GPCR\_RHODOPSIN.  
DR PROSITE: PS00237; G-PROTEIN\_RECEP\_F1\_1; 1.  
DR PROSITE: PS50262; G-PROTEIN\_RECEP\_F1\_2; 1.  
DR PROSITE: PS00238; OPSIN; 1.  
KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
FT Phosphorylation; G-protein coupled receptor.  
FT DOMAIN 1 46  
FT TRANSMEM 47 71  
FT DOMAIN 1 (POTENTIAL).  
FT CYTOPLASMIC.  
FT TRANSMEM 84 108  
FT DOMAIN 2 (POTENTIAL).  
FT EXTRACELLULAR.  
FT TRANSMEM 109 123  
FT DOMAIN 3 (POTENTIAL).  
FT CYTOPLASMIC.  
FT TRANSMEM 144 162  
FT DOMAIN 4 (POTENTIAL).  
FT EXTRACELLULAR.  
FT TRANSMEM 163 186  
FT DOMAIN 5 (POTENTIAL).  
FT CYTOPLASMIC.  
FT TRANSMEM 211 238  
FT DOMAIN 6 (POTENTIAL).  
FT EXTRACELLULAR.  
FT TRANSMEM 239 274  
FT DOMAIN 7 (POTENTIAL).  
FT CYTOPLASMIC.  
FT TRANSMEM 299 306  
FT DOMAIN 8 (POTENTIAL).  
FT EXTRACELLULAR.  
FT TRANSMEM 307 331  
FT DOMAIN 9 (POTENTIAL).  
FT CYTOPLASMIC.  
FT DISULFID 120 197  
FT BINDING 318 318  
FT CARBOHYD 17 17  
FT CARBOHYD 193 193  
SQ SEQUENCE 376 AA; 42139 MM; CCE401766AB06F26 CRC64;  
Query Match 39.4%; Score 43; DB 1; Length 376;  
Best Local Similarity 46.2%; Pred. No. 8.6;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Qy 2 PDINPAMYXXRGCI 14  
Db 40 PPMNPLWYSILGV 52  
RESULT 6

OPB2\_LIMPO STANDARD: PRT: 376 AA.  
 ID OPB2\_LIMPO  
 AC P35361;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ocellar opsin.  
 OS Limulus polyphemus (Atlantic horseshoe crab).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
 OC Limulidae; Limulus.  
 OX NCBI\_Taxid=6850;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Median ocelli;  
 RA MEDLINE=93317641; PubMed=8327495;  
 RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;  
 RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
 Limulus polyphemus";  
 RT Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
 CC IMMEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
 CC LINKED TO CIS-RETINAL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: OCELLAR CELLS; MEDIAN OCELLI.  
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
 CC BE PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 530 NM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC OPSIN SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: L03792; AAA28274.1; -  
 CC EMBL: L03782; AAA02499.1; -  
 CC PIR: A48197; A48197.  
 CC GCRD: GCR\_0585; -  
 CC InterPro: IPR000276; GPCR\_Rhodopsn.  
 CC InterPro: IPR001760; Opsin.  
 CC Pfam: PF00001; 7tm\_1; 1.  
 CC PRINTS: PR00237; GPCR\_Rhodopsn.  
 CC PROSITE: PS00237; G\_PROTEIN\_RECPT\_F1\_1; 1.  
 CC PROSITE: PS00262; G\_PROTEIN\_RECPT\_F1\_2; 1.  
 CC PROSITE: PS00238; OPSIN: 1.  
 CC Photoceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
 CC Phosphorylation; G-protein coupled receptor.  
 CC -----  
 CC DOMAIN 1 46  
 CC TRANSMEM 47 71  
 CC DOMAIN 72 83  
 CC TRANSMEM 84 108  
 CC DOMAIN 109 123  
 CC TRANSMEM 124 143  
 CC DOMAIN 144 162  
 CC TRANSMEM 163 186  
 CC DOMAIN 187 210  
 CC TRANSMEM 211 238  
 CC DOMAIN 239 274  
 CC TRANSMEM 275 298  
 CC DOMAIN 299 306  
 CC TRANSMEM 307 331  
 CC DOMAIN 332 376  
 CC DISULFID 120 197  
 CC BINDING 318 318  
 CC CARBOHYD 17 17  
 CC CARBOHYD 193 193  
 CC SEQUENCE 376 AA; 42111 MW; FA9647C40531CBF8 CRC64;

Query Match 39.4%; Score 43; DB 1; Length 376;  
 Best Local Similarity 46.2%; Pred. No. 8.6;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 PDINPAMYXXRG1 14  
 DB 40 PPMNPLMYSILGV 52  
 RESULT 7  
 CYCR\_CHRV1  
 ID CYCR\_CHRV1 STANDARD: PRT: 383 AA.  
 AC 082947;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Photosynthetic reaction center cytochrome C subunit precursor.  
 GN PufC.  
 OS Chromatium vinosum.  
 OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
 OC Allochromatium.  
 OX NCBI\_Taxid=1049;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-D;  
 RA Corson G.E., Nagashima K.V., Matsura K., Sakuragi Y., Ruwanthi W.,  
 RA Qin H., Allen R., Knaf D.B.;  
 RT "Primary structure of genes encoding light-harvesting and reaction  
 RT center proteins from Chromatium vinosum."  
 RT Submitted (MAR-1998) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A  
 CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO  
 CC OXIDIZED PRIMARY ELECTRON DONOR.  
 CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID  
 CC ANCHOR (BY SIMILARITY).  
 CC -1- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.  
 CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER  
 CC CYTOCHROME C SUBUNITS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: AB011811; BAA32742.1; -  
 CC HSSP: P07173; 6PRC.  
 CC InterPro: IPR003158; Cytc\_RC.  
 CC InterPro: IPR000345; Cytc\_heme\_bind.  
 CC Pfam: PF02276; CytoC\_RC; 1.  
 CC PROSITE: PS00190; CYTOCHROME\_C; 4.  
 CC Electron transport; Photosynthesis; Reaction center; Heme;  
 CC Membrane; Lipoprotein; Duplication; Signal.  
 CC -----  
 CC SIGNAL 1 22  
 CC CHAIN 23 383  
 CC LIPID 23 23  
 CC BINDING 107 107  
 CC BINDING 110 110  
 CC METAL 111 111  
 CC BINDING 152 152  
 CC BINDING 155 155  
 CC METAL 156 156  
 CC BINDING 247 247  
 CC BINDING 250 250  
 CC METAL 251 251  
 CC BINDING 307 307  
 CC BINDING 310 310



FT	METAL	311	311		IRON 4 (HEME AXIAL LIGAND)
FT					(BY SIMILARITY).
SO	SEQUENCE	383 AA;	41522 MW;	96BCD9FF1B9AE7E CRC64;	
	Query Match	Best Local Similarity	39.4%; Score 43; DB 1; Length 383;		
	Matches	7; Conservative	3; Mismatches	9; Indels	0; Gaps
OY	1	TPDINPANYXXRGIRVGR	19		
			:		
Db	263	TPORTTAVYAIRHVDING	281		
RESULT	8				
ID	YAAY_ECOLI	STANDARD:	PRT:	476 AA.	
AC	P30143:				
DT	01-JUL-1993 (Rel. 26, Created)				
DT	01-JUL-1993 (Rel. 26, Last sequence update)				
DT	16-OCT-2001 (Rel. 40, Last annotation update)				
DE	Putative transporter yaay.				
GN	YAAY OR B0007.				
OS	Escherichia coli.				
OC	Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;				
CC	Escherichia.				
OX	NCBI_TaxID=562;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-K12:				
RX	MEDLINE=92334977; PubMed=1630901;				
RA	Yura T., Mori H., Nagai H., Negata T., Ishihama A., Fujita N.,				
RA	Isono K., Michouchi K., Nakata A.;				
RT	"Systematic sequencing of the Escherichia coli genome: analysis of				
RL	the 0-2.4 min region."				
RL	Nucleic Acids Res. 20:3305-3308(1992).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	STRAIN-K12 / MG1655;				
RX	MEDLINE=97426617; PubMed=9278503;				
RA	Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,				
RA	Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,				
RA	Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,				
RA	Mau B., Siao Y.;				
RT	"The complete genome sequence of Escherichia coli K-12."				
RL	Science 277:1453-1474(1997).				
CC	-1 SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane				
CC	(potential).				
CC	-1 STABILITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY				
CC	(SAF). STRONG, TO H.INFLUENZAE H10183.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).				
CC	-----				
DR	EMBL; D10483; -; NOT_ANNOTATED_CDS.				
DR	EMBL; AE000111; AAC73118.1; -.				
DR	EcoGene; EG11555; yaay.				
DR	InferPro; IPR002293; AA_rel-permease_1.				
DR	InferPro; IPR001463; Na_ala_symp.				
DR	Pfam; PF01235; Na_Ala_symp; 1.				
DR	PRINTS; PR00175; NAALASPORT.				
DR	PROSITE; PS00873; NA_ALANINE_SYMP; 1.				
KW	Hypothetical protein; Transmembrane; Inner membrane; Transport;				
KW	Symport; Complete proteome.				
FT	TRANSMEM	4	24	POTENTIAL.	
FT	TRANSMEM	81	101	POTENTIAL.	
FT	TRANSMEM	141	161	POTENTIAL.	
FT	TRANSMEM	174	194	POTENTIAL.	

FT	TRANSMEM	207	227	POTENTIAL.			
FT	TRANSMEM	233	253	POTENTIAL.			
FT	TRANSMEM	300	320	POTENTIAL.			
FT	TRANSMEM	351	371	POTENTIAL.			
FT	TRANSMEM	391	411	POTENTIAL.			
FT	TRANSMEM	414	434	POTENTIAL.			
SO	SEQUENCE	476 AA;	51662 MW;	2F6EB2E12E126E63 CRC64;			
Query Match		39.4%;	Score 43;	DB 1;			
Best Local Similarity		44.4%;	Pred. No. 11;				
Matches	8;	Conservative	2;	Mismatches 2; Indels 6; Gaps 1;			
OY	3	DIN-----PAMVXXRGT	14				
		1:1	1111	11:			
DB	120	DVNGQFRCGPGAPVYMARGL	137				
RESULT	9						
NRPI_YEAST	STANDARD;	PRT;	719	AA.			
ID	NRPI_YEAST						
AC	P32770; Q12228;						
DT	01-OCT-1993 (Rel. 37, Created)						
DT	01-NOV-1997 (Rel. 35, Last sequence update)						
DT	01-MAR-2002 (Rel. 41, Last annotation update)						
DE	Asparagine-rich protein (ARP protein).						
GN	NRPI OR ARPI OR ARP OR YDL167C.						
OS	Saccharomyces cerevisiae (Baker's yeast).						
OC	Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;						
OC	Saccharomycetales; Saccharomycetaceae; Saccharomycetes.						
OX	NCBT_TaxID=4932;						
RM	[1]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN-AH22;						
RX	MEDLINE=93247548; PubMed=8483449;						
RT	Wehner E.P., Rao E., Brendel M.;						
RT	"Molecular structure and genetic regulation of SFA, a gene						
RT	responsible for resistance to formaldehyde in Saccharomyces						
RT	cerevisiae, and characterization of its protein product.";						
RL	Mol. Gen. Genet. 237:351-358(1993).						
RM	[2]						
RP	SEQUENCE FROM N.A.						
RC	STRAIN-S288C;						
RA	Pohl T.M.;						
RL	Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.						
CC	-1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.						
CC	-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).						
CC	-----						
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration						
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -						
CC	the European Bioinformatics Institute. There are no restrictions on its						
CC	use by non-profit institutions as long as its content is in no way						
CC	modified and this statement is not removed. Usage is and for commercial						
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>						
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).						
CC	-----						
DR	EMBL; X68020; CAA48159.1; -;						
DR	EMBL; Z67750; CAA91579.1; -;						
DR	EMBL; Z74215; CAA98741.1; -;						
DR	PIR; S31139; S31139.						
DR	HSSP; P04170; GRXN.						
DR	SCD; S0002326; NRPI.						
DR	InterPro: IPR000504; RRM.						
DR	InterPro: IPR001876; Znf-RanBP.						
DR	Pfam; PF00076; 1rm; 1.						
DR	Pfam; PF00641; zf-RanBP; 2.						
DR	SMART; SM00360; RRM; 1.						
DR	SMART; SM00547; Znf_RB2; 2.						
DR	PROSITE; PS00302; RRM; 1.						
DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.						
DR	PROSITE; PS03358; ZF_RANBP2_1; 2.						
DR	PROSITE; PS0199; ZF_RANBP2_2.						
KW	Nuclear protein; Zinc-finger; RNA-binding; Repeat.						

FT DOMAIN 226 322 RNA-BINDING (RRM).  
FT ZN\_FING 355 384 RANBP2-TYPE 1.  
FT ZN\_FING 581 610 RANBP2-TYPE 2.  
FT DOMAIN 490 564 ASN-RICH.  
FT CONFLICT 493 493 I -> N (IN REF. 1).  
SQ SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582669 CRC64;

Query Match 39.4%; Score 43; DB 1; Length 719;  
Best Local Similarity 50.0%; Pred. No. 16;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Oy 7 AWYXXRGIRPVG 18  
DB 244 SWFTQYGVPRVG 255

RESULT 10  
EXTL\_STRCO STANDARD; PRT; 402 AA.

AC Q9FBM3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
GN XSEA OR SCR7.29C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;

RN SEQUENCE FROM N.A.

RA Seeger K.J., Harris D., Cerdano A.M., Parkhill J., Barrell B.G.,  
Rajandream M.A.;

RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).

CC -1- CATALYTIC ACTIVITY: Exonuclease activity cleavage in either 5'- to 3'-  
or 3'- to 5'-direction to yield 5'-phosphonucleotides.

CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).

CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: AL391754; CAC05901.1; -  
DR InterPro: IPR003753; Exonuc\_VII\_L.  
DR InterPro: IPR002309; tRNA-synt\_2.  
DR Pfam: PF02601; Exonuc\_VII\_L; 1.  
DR Pfam: PF01336; tRNA-anti\_L; 1.  
KM Hydrolase; Nuclease; Exonuclease.  
SQ SEQUENCE 402 AA; 43882 MW; 145929A8372BAE08 CRC64;

Query Match 38.5%; Score 42; DB 1; Length 402;  
Best Local Similarity 42.9%; Pred. No. 13;  
Matches 9; Conservative 1; Mismatches 3; Indels 0; Gaps 1;

Oy 6 PANYXXRG-----IRPVG 18  
DB 89 PEWYAPRGOLSLRAEIKPVG 109

RESULT 11  
ARCC\_BACCHD STANDARD; PRT; 345 AA.

AC Q9K8V2;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-  
acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).  
GN ARCC OR BH2900.  
OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Scaphylococcus group; Bacillus.

OX NCBI\_TaxID=86665;

RN SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(++)  
+ phosphate = N-acetyl-5-glutamyl phosphate + NADPH.

CC -1- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.

CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC EMBL: AP01517; BAB0619.1; -  
DR InterPro: IPR000706; AGPR\_act\_site.

DR InterPro: IPR000534; Semialdh\_dh.

DR Pfam: PF01118; Semialdehyde\_dh; 1.

DR Pfam: PF02774; Semialdehyde\_dhc; 1.

DR Prodom: PD003765; AGPR\_act\_site; 1.

DR PROSITE: PS01224; ARCC; 1.  
KM Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.  
FW ACT SITE 149 149 BY SIMILARITY.

KT ACT SITE 149 149 BY SIMILARITY.

SQ SEQUENCE 345 AA; 38188 MW; 3E9F45DD09FC68EA CRC64;

Oy 2 PDINPANYXXRGIRPVG 18  
DB 106 PDVYEAMT-KROAAPVG 121

RESULT 12  
Y762\_METUA STANDARD; PRT; 342 AA.

AC Q58172;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein M0762.

GN M0762.

OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;

OC Methanococcus.  
OX NCBI\_TaxID=2190;

RN SEQUENCE FROM N.A.

RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RA MEDLINE-96337999; PubMed-8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
 RA Klenk H.-P., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Clontz H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii".  
 RL Science 273:1058-1073(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: STRONG, TO M.JANNASCHII M00576 AND TO S.POMBE MALATE  
 PERMEASE (MAE1).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U67521; AAB98753.1; -.  
 DR TIGR; M00762; -.  
 KW Hypothetical protein; Transmembrane; Transport; Complete proteome.  
 FT TRANSMEM 8 28 POTENTIAL.  
 FT TRANSMEM 39 59 POTENTIAL.  
 FT TRANSMEM 79 99 POTENTIAL.  
 FT TRANSMEM 108 128 POTENTIAL.  
 FT TRANSMEM 142 162 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 FT TRANSMEM 207 227 POTENTIAL.  
 FT TRANSMEM 242 262 POTENTIAL.  
 FT TRANSMEM 276 296 POTENTIAL.  
 FT TRANSMEM 304 324 POTENTIAL.  
 SQ SEQUENCE 342 AA; 39534 MW; 08EFC3E2C4955D8 CRC64;

Query Match 37.6%; Score 41; DB 1; Length 342;  
 Best Local Similarity 53.3%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 4 INPAMYXXRGIRPYG 18  
 Db 139 VNPQWY----IPYVG 149

RESULT 13  
 Y576\_METJA STANDARD; PRT; 347 AA.  
 AC 057996;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein M00576.  
 OS M00576.  
 GN Methanococcus jannaschii.  
 OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
 OC Methanococcus.  
 NCBI\_TaxID=2190;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
 RA MEDLINE-96337999; PubMed-8688087;  
 RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
 RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
 RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
 RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
 RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhrman J.L., Nguyen D.,  
 RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,

RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
 RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
 RT "Complete genome sequence of the methanogenic archaeon, Methanococcus  
 jannaschii".  
 RL Science 273:1058-1073(1996).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: STRONG, TO M.JANNASCHII M00762 AND TO S.POMBE MALATE  
 PERMEASE (MAE1).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; U67506; AAB98567.1; -.  
 DR TIGR; M00576; -.  
 KW Hypothetical protein; Transmembrane; Transport; Complete proteome.  
 FT TRANSMEM 15 35 POTENTIAL.  
 FT TRANSMEM 46 66 POTENTIAL.  
 FT TRANSMEM 84 104 POTENTIAL.  
 FT TRANSMEM 111 131 POTENTIAL.  
 FT TRANSMEM 149 169 POTENTIAL.  
 FT TRANSMEM 182 202 POTENTIAL.  
 FT TRANSMEM 214 234 POTENTIAL.  
 FT TRANSMEM 249 269 POTENTIAL.  
 FT TRANSMEM 283 303 POTENTIAL.  
 FT TRANSMEM 312 332 POTENTIAL.  
 SQ SEQUENCE 347 AA; 39556 MW; 632F671A31DE183 CRC64;

Query Match 37.6%; Score 41; DB 1; Length 347;  
 Best Local Similarity 53.3%; Pred. No. 17;  
 Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

QY 4 INPAMYXXRGIRPYG 18  
 Db 146 VNPQWY----IPYVG 156

RESULT 14  
 CYGR\_ARBPV STANDARD; PRT; 986 AA.  
 AC P1528;  
 DT 01-OCT-1989 (Rel. 12, Created)  
 DT 01-OCT-1989 (Rel. 12, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Reasact receptor precursor (Guanylate cyclase) (EC 4.6.1.2).  
 OS Arabacia punctulata (Punctate sea urchin).  
 OC Eukaryota; Metazoa; Echinodermata; Eleutherozoa; Echinozoa;  
 OC Echinoidea; Euechinoidea; Echinacea; Arbacoidea; Arbacoidea; Arbacoidea;  
 OC NCBI\_TaxID=7641;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RC TISSUE-Testis;  
 RA MEDLINE-88318927; PubMed-2901039;  
 RA Singh S., Lowe D.G., Thorpe D.S., Rodriguez H., Kuang W.-J.,  
 RA Dangott L.J., Chinkers M., Goeddel D.V., Garbers D.L.;  
 RT "Membrane guanylate cyclase is a cell-surface receptor with homology  
 RT to protein kinases.";  
 RL Nature 334:708-712(1988).  
 CC -1- FUNCTION: IMPLICATED AS A CELL-SURFACE RECEPTOR ON SPERMATOZOA  
 CC FOR 'REASACT' A CHEMOTACTIC PEPTIDE, AND ON VARIOUS OTHER CELLS  
 CC AS A RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE.  
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X12874; CAA1367.1; -  
 DR PIR: S05480; OYURGA.  
 DR InterPro: IPR001828; ANF\_receptor.  
 DR InterPro: IPR000719; Euk\_kinase.  
 DR InterPro: IPR001054; Guanylyl\_cyclase.  
 DR Pfam: PF01094; ANF\_receptor.1.  
 DR Pfam: PF00211; guanylate\_cyc.1.  
 DR Pfam: PF00069; kinase.2.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 DR Receptor: Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
 KW CGMP synthetasis; Signal.  
 FT SIGNAL 1 21  
 FT CHAIN 1 986 RESACT RECEPTOR  
 FT DOMAIN 22 507 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 508 528 POTENTIAL.  
 FT 529 986 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 568 836 PROTEIN KINASE LIKE.  
 FT CARBOHYD 185 185 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 361 361 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 410 410 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SQ SEQUENCE 986 AA; 111284 MW; B40238A74CACAFC52 CRC64;

Query Match 37.6%; Score 41; DB 1; Length 986;  
 Best Local Similarity 62.5%; Pred. No. 46;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 2 PDINPAMY 9  
 DB 475 PDLNPMWH 482

RESULT 15  
 Y493\_MYCTU  
 ID Y493\_MYCTU STANDARD: PRT: 329 AA.  
 AC Q11158;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 35.4 kDa protein RV0493C.  
 GN RV0493C OR MT0513 OR MTCY2069.19C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
 Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 Ruter S., Seeger K., Skelton S., Squares S., Squares R.,  
 Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CDC 1551 / Oshkosh;  
 RA Fleischmann R.D., Alland D., Eisen J.A., Carpenter L., White O.,  
 Peterson J., DeBoy R., Dodson R., Gwinn M.L., Haft D., Hickey E.,  
 Kolonay J.F., Nelson W.C., Umayam L.A., Ermolaeva M.D., Salzberg S.L.,  
 Delcher A., Utterback T., Feldman J., Khouri H., Gill J., Mikula A.,  
 Bishai W.;

RT "Whole genome comparison of Mycobacterium tuberculosis clinical and  
 RT laboratory strains.";  
 RL Submitted (Apr-2001) to the EMBL/GenBank/DBJ databases.  
 CC -1- SIMILARITY: SOME, TO M LEPRAE B2168\_F2.93.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: Z77162; CAB00954.1; -  
 DR EMBL: AE006952; AAK44736.1; -  
 DR TIGR: MT0513; -  
 DR TubercuList: RV0493C; -  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 329 AA; 35427 MW; BE4B524750277B56 CRC64;

Query Match 36.7%; Score 40; DB 1; Length 329;  
 Best Local Similarity 42.9%; Pred. No. 24;  
 Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;  
 QY 2 PDINPAMYXXRGIR 15  
 DB 90 PAAGPAMFDIAGVR 103

Search completed: September 13, 2002, 09:30:43  
 Job time: 1133 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:57 ; Search time 136.62 Seconds  
(without alignments)  
1.251 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_16\_22  
Perfect score: 30  
Sequence: 1 PVGRPX 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 2442594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/prodata/2/1aa/5a\_COMB.pep:\*
- 2: /cgn2\_6/prodata/2/1aa/5b\_COMB.pep:\*
- 3: /cgn2\_6/prodata/2/1aa/6a\_COMB.pep:\*
- 4: /cgn2\_6/prodata/2/1aa/6b\_COMB.pep:\*
- 5: /cgn2\_6/prodata/2/1aa/PCOTUS\_COMB.pep:\*
- 6: /cgn2\_6/prodata/2/1aa/beckfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	93.3	15	4	US-08-776-971-93
2	28	93.3	20	3	US-09-105-678A-34
3	28	93.3	20	3	US-09-105-678A-40
4	28	93.3	20	3	US-09-105-678A-46
5	28	93.3	20	4	US-08-776-971-8
6	28	93.3	20	4	US-08-776-971-50
7	28	93.3	20	4	US-08-776-971-64
8	28	93.3	20	4	US-08-776-971-98
9	28	93.3	20	4	US-09-421-208-34
10	28	93.3	20	4	US-09-421-208-40
11	28	93.3	20	4	US-09-421-208-46
12	28	93.3	21	3	US-09-105-678A-28
13	28	93.3	21	3	US-09-105-678A-35
14	28	93.3	21	3	US-09-105-678A-41
15	28	93.3	21	3	US-09-105-678A-47
16	28	93.3	21	4	US-08-776-971-9
17	28	93.3	21	4	US-08-776-971-51
18	28	93.3	21	4	US-08-776-971-65
19	28	93.3	21	4	US-09-421-208-28
20	28	93.3	21	4	US-09-421-208-35
21	28	93.3	21	4	US-09-421-208-41
22	28	93.3	21	4	US-09-421-208-47
23	28	93.3	22	3	US-09-105-678A-36
24	28	93.3	22	3	US-09-105-678A-42
25	28	93.3	22	3	US-09-105-678A-48
26	28	93.3	22	4	US-08-776-971-10
27	28	93.3	22	4	US-08-776-971-52

28	28	93.3	22	4	US-08-776-971-66	Sequence 66, Appl
29	28	93.3	22	4	US-08-776-971-73	Sequence 73, Appl
30	28	93.3	22	4	US-09-421-208-36	Sequence 36, Appl
31	28	93.3	22	4	US-09-421-208-42	Sequence 42, Appl
32	28	93.3	22	4	US-09-421-208-48	Sequence 48, Appl
33	28	93.3	31	3	US-09-105-678A-7	Sequence 7, Appl
34	28	93.3	31	3	US-09-105-678A-8	Sequence 8, Appl
35	28	93.3	31	3	US-09-105-678A-9	Sequence 9, Appl
36	28	93.3	31	3	US-09-105-678A-31	Sequence 31, Appl
37	28	93.3	31	3	US-09-105-678A-37	Sequence 37, Appl
38	28	93.3	31	3	US-09-105-678A-43	Sequence 43, Appl
39	28	93.3	31	4	US-09-172-353-4	Sequence 43, Appl
40	28	93.3	31	4	US-08-776-971-5	Sequence 4, Appl
41	28	93.3	31	4	US-08-776-971-47	Sequence 47, Appl
42	28	93.3	31	4	US-08-776-971-61	Sequence 61, Appl
43	28	93.3	31	4	US-08-776-971-97	Sequence 97, Appl
44	28	93.3	31	4	US-09-421-208-7	Sequence 7, Appl
45	28	93.3	31	4	US-09-421-208-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-776-971-93  
Sequence 93, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93  
US-08-776-971-93

Query Match 93.3%; Score 28; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 2.1;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 11 PVGRF 15

## RESULT 2

US-09-105-678A-34  
Sequence 34, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 93.3%; Score 28; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 16 PVGRF 20

## RESULT 3

US-09-105-678A-40  
Sequence 40, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 93.3%; Score 28; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 16 PVGRF 20

## RESULT 4

US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 93.3%; Score 28; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVRGF 5  
|||||  
DB 16 PVRGF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habeta, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVRGF 5  
|||||  
DB 16 PVRGF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habeta, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 16 PVGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 16 PVGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400



TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
US-08-776-971-98  
SEQUENCE DESCRIPTION: SEQ ID NO: 98:

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVRGF 5  
16 PVRGF 20

RESULT 9  
US-09-421-208-34  
; Sequence 34, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 PVRGF 5

Db 16 PVRGF 20

RESULT 10  
US-09-421-208-40  
; Sequence 40, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVRGF 5  
16 PVRGF 20

RESULT 11  
US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide.  
US-09-421-208-46

Query Match 93.3%; Score 28; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 2.8;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 16 PVGRF 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product= "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product= "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product= "Gly-OH or Gly-Arg"  
US-09-105-678A-28

Query Match 93.3%; Score 28; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 16 PVGRF 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: FLOPPY disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 93.3%; Score 28; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
DB 16 PVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 93.3%; Score 28; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 PVGRF 5  
DB 16 PVGRF 20

DB 16 PVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 93.3%; Score 28; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 2.9;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 PVGRF 5  
DB 16 PVGRF 20

Search completed: September 13, 2002, 09:20:57  
Job time: 622 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 13, 2002, 09:23:58 : Search time 172.41 Seconds  
(without alignments)  
3.901 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_16\_22

Perfect score: 30

Sequence: 1 PVGRFXX 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR71:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	93.3	83	2 JC7607	prolactin-releasin
2	28	93.3	154	2 T34825	hypothetical prote
3	28	93.3	170	2 S76067	hypothetical prote
4	28	93.3	173	2 S76779	hypothetical prote
5	28	93.3	193	2 T44148	hypothetical prote
6	28	93.3	246	2 H98292	3-oxoacyl-(acyl-ca
7	28	93.3	246	2 A12990	3-oxoacyl-(acyl-ca
8	28	93.3	251	2 AB3392	phosphotyrosyl pho
9	28	93.3	257	2 F83104	probable enoyl-CoA
10	28	93.3	260	2 T36846	probable dehydroge
11	28	93.3	294	2 F84922	hypothetical prote
12	28	93.3	299	2 C70643	hypothetical prote
13	28	93.3	300	1 S08244	conserved hypothet
14	28	93.3	367	2 T44687	cobalamn biosynth
15	28	93.3	424	2 D69399	3-ketoacyl-CoA thl
16	28	93.3	433	2 B47041	hypothetical prote
17	28	93.3	445	2 T35893	FAD-dependent oxid
18	28	93.3	457	2 AG2728	polysaccharide blo
19	28	93.3	458	2 AC2935	conserved hypothet
20	28	93.3	497	2 D98347	hypothetical prote
21	28	93.3	501	2 S66763	hypothetical prote
22	28	93.3	511	2 T40334	hypothetical prote
23	28	93.3	522	2 H97509	probable glycosyl
24	28	93.3	559	2 G83897	uroconase (urocana
25	28	93.3	585	2 C69336	probable electron
26	28	93.3	601	2 H69274	probable long-chain
27	28	93.3	612	2 T36210	conserved hypothet
28	28	93.3	847	2 F64847	glucan biosynthesi
29	28	93.3	847	2 G85666	hypothetical prote

30	28	93.3	847	2 AH0636	periplasmic glucan
31	28	93.3	857	2 C90807	membrane glycosyl
32	28	93.3	1057	1 OYRFR	atrial natriuretic
33	28	93.3	1057	2 I57963	natriuretic peptid
34	28	93.3	1057	2 I55319	guanylyl cyclase A
35	28	93.3	1051	1 OYHAR	natriuretic peptid
36	28	93.3	2476	2 T34022	zonadhesin - pig
37	27	90.0	140	2 S04165	hypothetical prote
38	27	90.0	159	2 S14179	hypothetical prote
39	27	90.0	159	2 S57630	hypothetical prote
40	27	90.0	168	2 F95365	NADH dehydrogenase
41	27	90.0	199	2 AH3320	hypothetical prote
42	27	90.0	239	2 AB1887	MoA/ToIq/Exb pro
43	27	90.0	261	2 D98337	L-Iditolisorbitol d
44	27	90.0	261	2 AF2945	dehydrogenase Atu3
45	27	90.0	308	2 AC1053	probable membrane

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; M0ID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 93.3%; Score 28; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
Db 48 PVGRF 52

RESULT 2  
T34825  
hypothetical protein SC2E9.14 SC2E9.14 - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34825  
R:Oliver, K.; Harris, D.; Parkhill, J.; Bartrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, January 1998  
A:Reference number: Z11558  
A:Accession: T34825  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-154 <OLIS>  
A:Cross-references: EMBL:AL021530; PTDN:CAA16482.1; GSPDB:GN00070; SCODEB:SC2E9.14  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCODEB:SC2E9.14

Query Match 93.3%; Score 28; DB 2; Length 154;  
Best Local Similarity 100.0%; Pred. No. 36;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||

DB 32 PVGRF 36

## RESULT 3

S76067  
Hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S76067

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76067

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-170 <KAN>

A:Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BA010045.1; PID:d101065

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

## Query Match

Best Local Similarity 93.3%; Score 28; DB 2; Length 170;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||

DB 47 PVGRF 51

## RESULT 4

S76779  
Hypothetical protein - *Synechocystis* sp. (strain PCC 6803)

C:Species: *Synechocystis* sp.

A:Variety: PCC 6803

C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999

C:Accession: S76779

R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima, N.;

O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yasuda

DNA Res. 3, 109-136, 1996

A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocystis*

s.

A:Reference number: S74322; MUID:97061201

A:Accession: S76779

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-173 <KAN>

A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BA018691.1; PID:d101942

A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

## Query Match

Best Local Similarity 93.3%; Score 28; DB 2; Length 173;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||

DB 95 PVGRF 99

## RESULT 5

T44148

hypothetical protein B4 [imported] - human herpesvirus 6 (strain Z29)

C:Species: human herpesvirus 6

A:Variety: strain Z29

C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 02-Jun-2000

C:Accession: T44148  
R:Domínguez, G.; Dambaugh, T. R.; Stamey, F. R.; Dewhurst, S.; Inoue, N.; Pellett, P. E.  
J. Virol. 73, 8040-8052, 1999

A:Title: Human herpesvirus 6B genome sequence: coding content and comparison with hum

A:Reference number: 222734; MUID:99412318

A:Accession: T44148

A:Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-193 <DOM>

A:Cross-references: EMBL:AF157706; PIDN:AA049620.1

A:Experimental source: strain Z29; variant B

C:Genetics:

A:Note: B4

## Query Match

Best Local Similarity 93.3%; Score 28; DB 2; Length 193;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||

DB 106 PVGRF 110

## RESULT 6

H98292  
3-oxoacyl-(acyl-carrier-protein) reductase VC2021 [imported] - *Agrobacterium tumefaci*

C:Species: *Agrobacterium tumefaciens*

C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002

C:Accession: H98292

R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouellet, B.; Goldm

A.; Liu, F.; Woliam, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Markelz,

Science 294, 2323-2328, 2001

A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent *Agrobacterium*

A:Reference number: A97359; PMID:11743194

A:Accession: H98292

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-246 <KOR>

A:Cross-references: GB:AE007870; PIDN:AAK89866.1; PID:g15159809; GSPDB:GN00170

C:Genetics:

A:Gene: AGR\_L\_2601

A:Map position: linear chromosome

C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

## Query Match

Best Local Similarity 93.3%; Score 28; DB 2; Length 246;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||

DB 207 PVGRF 211

## RESULT 7

A12990  
3-oxoacyl-(acyl-carrier-protein) reductase [imported] - *Agrobacterium tumefaciens* (st

C:Species: *Agrobacterium tumefaciens*

C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 01-Feb-2002

C:Accession: A12990

R:Wood, D. W.; Sedubal, J. C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G. E.; Chen, Y.; Woo

erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; MCCI

Science 294, 2317-2323, 2001

A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam

ster, E. W.

A:Title: The Genome of the Natural Genetic Engineer *Agrobacterium tumefaciens* C58.

A:Reference number: AB2577; PMID:11743193

A:Accession: A12990

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-246 <KOR>

A:Cross-references: GB:AE008689; PIDN:AAL44343.1; PID:g17741935; GSPDB:GN00187  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: fabG  
A:Map position: linear chromosome  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 93.3%; Score 28; DB 2; Length 246;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 207 PVGRF 211

RESULT 8  
AB3392  
phosphoryl phosphatase activator (ptpa) [imported] - Brucella melitensis (strain 16M)  
C:Species: Brucella melitensis  
C>Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 15-Feb-2002  
C:Accession: AB3392  
R:Delvecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen Brucella melitens  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AB3392  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-251 <NR>  
A:Cross-references: GB:AE008917; PIDN:AAL52301.1; PID:g17983092; GSPDB:GN00190  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BME1120  
A:Map position: 1  
C:Superfamily: 2,4-dihydroxyhept-2-ene-1,7

Query Match 93.3%; Score 28; DB 2; Length 251;  
Best Local Similarity 100.0%; Pred. No. 60;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 68 PVGRF 72

RESULT 9  
F83104  
probable enoyl-CoA hydratase/isomerase PA4330 [imported] - Pseudomonas aeruginosa (strai  
C:Species: Pseudomonas aeruginosa  
C>Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 31-Dec-2000  
C:Accession: F83104  
R:Stover, C.K.; Pham, X.Q.; Ewitt, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llm,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pat  
A:Reference number: AB2950; MUID:20437337  
A:Accession: F83104  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-257 <STO>  
A:Cross-references: GB:AE004849; GB:AE004091; MID:g9950550; PIDN:AA607718.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:  
A:Gene: PA4330  
C:Superfamily: naphthoate synthase; enoyl-CoA hydratase homology

Query Match 93.3%; Score 28; DB 2; Length 257;

Best Local Similarity 100.0%; Pred. No. 61;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 PVGRF 5  
|||||  
DB 80 PVGRF 84

RESULT 10  
T36846  
probable dehydrogenase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 11-Jan-2000  
C:Accession: T36846  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; RaJandream, M.A.  
submitted to the EMBL Data Library, September 1998  
A:Reference number: Z21615  
A:Accession: T36846  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-260 <OLI>  
A:Cross-references: EMBL:AL031541; PIDN:CAA20822.1; GSPDB:GN00070; SCODEB:SCI35.33C  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCODEB:SCI35.33c  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 93.3%; Score 28; DB 2; Length 260;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 216 PVGRF 220

RESULT 11  
F84922  
hypothetical protein At2g48040 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 02-Feb-2001  
C:Accession: F84922  
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Beaulieu, M.L.; Town, C.D.; Fujii, C.Y.  
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon,  
euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter  
Nature 402, 761-768, 1999  
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.  
A:Reference number: AB4420; MUID:2083487  
A:Accession: F84922  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-294 <STO>  
A:Cross-references: GB:AE002093; MID:g4249411; PIDN:AA013708.1; GSPDB:GN00139  
C:Genetics:  
A:Gene: At2g48040  
A:Map position: 2

Query Match 93.3%; Score 28; DB 2; Length 294;  
Best Local Similarity 100.0%; Pred. No. 70;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 221 PVGRF 225

RESULT 12  
C70643  
hypothetical protein Rv0712 - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999

C:Accession: C70643  
 R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fellwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
 A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
 A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
 A:Reference number: A70500; MUID:98295987  
 A:Accession: C70643  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1\*299 <COL>  
 A:Cross-references: GB:284395; GB:AL123456; NID:93261698; PIDN:CAB06436.1; PID:e293289;  
 A:Experimental source: strain H37Rv  
 C:Genetics:  
 A:Gene: RV0712

Query Match 93.3%; Score 28; DB 2; Length 299;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 199 PVGRF 203

## RESULT 13

S08244  
 conserved hypothetical protein MJ1225 homolog - Thermophilum pendens  
 N:Alternate names: hypothetical protein 1  
 C:Species: Thermophilum pendens  
 C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
 A:Accession: S08244  
 R:Kjems, J.; Leffers, H.; Olesen, T.; Ingelore, H.; Garrett, R.A. submitted to the EMBL Data Library, March 1989  
 A:Reference number: S08244  
 A:Accession: S08244  
 A:Molecule type: DNA  
 A:Residues: 1-300 <KJ>  
 A:Cross-references: EMBL:X14835; NID:948225; PID:948226  
 C:Superfamily: conserved hypothetical protein MJ1225; CBS homolog  
 C:Keywords: duplication  
 F:13-61/Domain: CBS homolog <CBS1>  
 F:92-140/Domain: CBS homolog <CBS2>  
 F:155-202/Domain: CBS homolog <CBS3>  
 F:230-277/Domain: CBS homolog <CBS4>

Query Match 93.3%; Score 28; DB 1; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 71;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 7 PVGRF 11

## RESULT 14

T44687  
 cobalamin biosynthesis protein [Imported] - Bacillus megaterium  
 C:Species: Bacillus megaterium  
 C:Date: 21-Jan-2000 #sequence\_revision 21-Jan-2000 #text\_change 21-Jul-2000  
 A:Accession: T44687  
 R:Raux, E.; Lanois, A.; Warren, M.J.; Rambach, A.; Thermes, C. Biochem. J. 335, 159-166, 1998  
 A:Title: Cobalamin (vitamin B12) biosynthesis: identification and characterization of a  
 A:Reference number: Z22829; MUID:98416126  
 A:Accession: T44687  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-367 <RAU>

A:Cross-references: EMBL:AJ000758; NID:93724036; PIDN:CAA04311.1; PID:93724042

A:Experimental source: strain DSM 509  
 C:Genetics:  
 A:Note: cblD  
 C:Superfamily: Methanobacterium cobalamin biosynthesis protein D

Query Match 93.3%; Score 28; DB 2; Length 367;  
 Best Local Similarity 100.0%; Pred. No. 88;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 46 PVGRF 50

## RESULT 15

D69399  
 3-ketoacyl-CoA thiolase (fadA-2) homolog - Archaeoglobus fulgidus  
 C:Species: Archaeoglobus fulgidus  
 C:Date: 05-Dec-1997 #sequence\_revision 05-Dec-1997 #text\_change 18-Jun-1999  
 A:Accession: D69399  
 R:Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dod  
 .; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E  
 Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.F.; McDonald, L.  
 Nature 390, 364-370, 1997  
 A:Authors: Overbeek, T.; Cotton, M.D.; Spriggs, T.; Arlisch, P.; Kaine, B.P.; Sykes,  
 Smith, H.O.; Woese, C.R.; Venter, J.C.  
 A:Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing arch  
 A:Reference number: A69250; MUID:98049343  
 A:Accession: D69399  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-424 <KLE>  
 A:Cross-references: GB:AE001021; GB:AE000782; NID:92689344; PIDN:AAB90044.1; PID:9264  
 C:Superfamily: acetyl-CoA acetyltransferase

Query Match 93.3%; Score 28; DB 2; Length 424;  
 Best Local Similarity 100.0%; Pred. No. 1e+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 11 PVGRF 15

Search completed: September 13, 2002, 09:23:58  
 Job time: 773 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:43 ; Search time 80.21 Seconds

(without alignments)  
3.379 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_16\_22

Perfect score: 30

Sequence: 1 PVGRFX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	93.3	83	1	PRRP_RAT
2	28	93.3	87	1	PRRP_HUMAN
3	28	93.3	98	1	PRRP_BOVIN
4	28	93.3	292	1	Y4AD_RHISN
5	28	93.3	300	1	YR33_THERPE
6	28	93.3	330	1	AR72_HUMAN
7	28	93.3	331	1	AR73_HUMAN
8	28	93.3	480	1	GAG1_RHIME
9	28	93.3	526	1	CATL_BRARE
10	28	93.3	847	1	MODH_ECOLI
11	28	93.3	1057	1	ANPA_MOUSE
12	28	93.3	1057	1	ANPA_RAT
13	28	93.3	1061	1	ANPA_HUMAN
14	28	93.3	2476	1	ZAN_PIG
15	27	90.0	140	1	YPL3_STANU
16	27	90.0	168	1	NDE2_RHIME
17	27	90.0	250	1	LINC_PSEPA
18	27	90.0	332	1	OPT_HUMAN
19	27	90.0	338	1	DCUP_AOUAE
20	27	90.0	400	1	PCAF_PSEPU
21	27	90.0	787	1	SP3E_BACSU
22	25	83.3	128	1	IGF2_CAVPO
23	25	83.3	129	1	IGF2_MUSVI
24	25	83.3	130	1	RL3_PIG
25	25	83.3	151	1	Y17K_BPPA
26	25	83.3	155	1	IGF2_BOVIN
27	25	83.3	156	1	IGF2_ABRPE
28	25	83.3	179	1	IGF2_SHEEP
29	25	83.3	180	1	IGF2_HUMAN
30	25	83.3	180	1	IGF2_MOUSE
31	25	83.3	180	1	IGF2_RAT
32	25	83.3	181	1	IGF2_PIG
33	25	83.3	184	1	PINE_ECOLI

34	25	83.3	186	1	DNIV_BPP1
35	25	83.3	186	1	DNIV_BPP7
36	25	83.3	193	1	DNIV_BPMU
37	25	83.3	215	1	PIMT_METJA
38	25	83.3	228	1	P213_HUMAN
39	25	83.3	246	1	FABG_THEMA
40	25	83.3	248	1	FABG_AOUAE
41	25	83.3	263	1	NDHM_NEUCR
42	25	83.3	285	1	Y048_HAEIN
43	25	83.3	325	1	YGAT_ECOLI
44	25	83.3	325	1	YGAT_SALTY
45	25	83.3	346	1	EP1A_ETIMO

#### ALIGNMENTS

RESULT 1

ID	PRRP_RAT	STANDARD	PRT	83 AA.
AC	P81278;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].			
DE	PRH.			
GN	Rattus norvegicus (Rat).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.			
OX	NCHI_TaxID=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=98268781; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,			
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,			
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RT	"A proactin-releasing peptide in the brain.";			
RL	Nature 393:272-276(1998).			
RN	[2]			
RP	TISSUE SPECIFICITY.			
RX	PubMed=10498338;			
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,			
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,			
RT	Sumino Y., Fujino M.;			
RL	"Tissue distribution of proactin-releasing peptide (PRP) and its receptor.";			
RL	Regul. Pept. 83:1-10(1999).			
CC	-I- FUNCTION: Stimulates proactin (PRL) release and regulates the expression of proactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.			
CC	-I- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sb.ch/announce/">http://www.isb-sb.ch/announce/</a> or send an email to <a href="mailto:license@isb-sb.ch">license@isb-sb.ch</a> ).			
CC	-----			
DR	EMBL: AB015418; BAA29026.1;			
KW	Hormone; Amidation; Signal; Cleavage on pair of basic residues.			
FT	SIGNAL	1	21	
FT	PEPTIDE	22	52	
FT	PEPTIDE	33	52	
FT	MOD. RES	52	52	
SO	SEQUENCE	83 AA;	9215 MW;	

Query Match 93.3%; Score 28; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 6.3;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 48 PVGRF 52

## RESULT 2

PRRP\_HUMAN

ID PRRP\_HUMAN STANDARD; PRT; 87 AA.

AC P01277;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].  
 GN PRH.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606.

RN [1]  
 RP SEQUENCE FROM N.A.

RC TISSUE=Brain;  
 RX MEDLINE=98268781; Pubmed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).

RN [2]  
 RP TISSUE SPECIFICITY.

RX Pubmed=10498338;

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RT "Issue distribution of prolactin-releasing peptide (Prp) and its receptor.";  
 RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB015419; BAA29027.1; -;  
 DR MIM; 602653; -;

KM Hormone; Amidation; Signal.

FT SIGNAL 1 22 BY SIMILARITY.  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP1.  
 FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP2.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).

SO SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

Query Match 93.3%; Score 28; DB 1; Length 87;

Best Local Similarity 100.0%; Pred. No. 6.6;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 49 PVGRF 53

RESULT 3  
 PRRP\_BOVIN  
 ID PRRP\_BOVIN STANDARD; PRT; 98 AA.

AC P01264;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].  
 GN PRH.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OX Bovidae; Bovine; Bos.  
 OX NCBI\_TaxID=9913;

RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RC TISSUE=Brain;  
 RX MEDLINE=98268781; Pubmed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AB015417; BAA29025.1; -;  
 KM Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP1.  
 FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP2.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).

SO SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 93.3%; Score 28; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 7.4;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
 |||||  
 DB 49 PVGRF 53

## RESULT 4

YAAD\_RHISN

ID YAAD\_RHISN STANDARD; PRT; 292 AA.

AC P55351;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 15-JUL-1998 (Rel. 36, Last annotation update)  
 DE Hypothetical 32.5 kDa protein YAAD.

GN YAAD.  
 OS Rhizobium sp. (strain NGR234).

OG Plasmid sym pNCR234a.

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Rhizobium.

OX NCBI\_TaxID=394;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE-97305956; PubMed-9163424;  
 RA Frelberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,  
 RA Perret X.;  
 RT "Molecular basis of symbiosis between Rhizobium and Legumes.";  
 RL Nature 387:394-401(1997).  
 CC -1- FUNCTION: PROBABLY PART OF AN OPERON YAAABD INVOLVED IN THE  
 CC SYNTHESIS OF AN ISOPRENOID COMPOUND.  
 CC -1- SIMILARITY: SOME, TO PHYTOENE AND SQUALENE SYNTHETASES.  
 CC -1- SIMILARITY: SOME, TO Y4AC.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, AF000064; AAB91602.1;  
 KW Hypothetical protein; Plasmid; Isoprene biosynthes.  
 SQ SEQUENCE 292 AA; 32539 MW; 1055E5E7F3614379 CRC64;

Query Match 93.3%; Score 28; DB 1; Length 292;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYGRF 5  
 DB 127 PYGRF 131

RESULT 5  
 YR33\_THEPE STANDARD: PRT; 300 AA.

AC 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical 33.4 kDa protein in ribosomal RNA operon.  
 OS Thermophilum pendens.  
 CC Archaea; Crenarchaeota; Thermoproteales; Thermofillicae; Thermophilum.  
 OX NCBI\_TaxID=2289;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-HVY3 / DSM 2475;  
 RA Klems J., Jeffers H., Olesen T., Ingelore H., Garrett R.A.;  
 RT "Sequence, organization and transcription of the ribosomal RNA operon  
 RT and the downstream RNA and protein genes in the archaeobacterium  
 RT Thermophilum pendens.";  
 RL Syst. Appl. Microbiol. 13:117-127(1990).  
 CC -1- SIMILARITY: CONTAINS 4 CBS DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL, X14835; CAA32944.1;  
 DR PIR, S08244; S08244.  
 DR InterPro: IPR000644; CBS.  
 DR Pfam: PF00571; CBS; 4.  
 DR SMART: SM00116; CBS; 4.  
 KM Hypothetical protein; Repeat; CBS domain.  
 FT DOMAIN 8 61 CBS 1.  
 FT DOMAIN 87 140 CBS 2.  
 FT DOMAIN 150 202 CBS 3.  
 FT DOMAIN 224 276 CBS 4.  
 SQ SEQUENCE 300 AA; 33437 MW; D811A313D37A4293 CRC64;

Query Match 93.3%; Score 28; DB 1; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 23;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PYGRF 5  
 DB 7 PYGRF 11

RESULT 6

AR72\_HUMAN STANDARD: PRT; 330 AA.  
 AC 043488; 075749;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Aflatoxin B1 aldehyde reductase 1 (EC 1.-.-.-) (AFB1-AR 1)  
 DE (Aldoketoreductase 7).  
 GN AKR7A2 OR AFAR OR AKR7.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE-98244807; PubMed-9576847;  
 RA Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;  
 RT "Molecular cloning, expression and catalytic activity of a human AKR7  
 RT member of the aldo-keto reductase superfamily: evidence that the  
 RT major 2-carboxybenzaldehyde reductase from human liver is a homologue  
 RT of rat aflatoxin B1-aldehyde reductase.";  
 RL Biochem. J. 332:21-34(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE-99040634; PubMed-9823300;  
 RA Pram C., Savelyeva L., Perri P., Schwab M.;  
 RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-  
 RT 1p36.1 in a region frequently altered in human tumor cells.";  
 RL Cancer Res. 58:5014-5018(1998).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Hall R.;  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.

-1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING  
 DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOL. COULD BE  
 INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC  
 EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE  
 REDUCTASE.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation-  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL, AF026947; AAC52104.1;  
 DR EMBL, Y16675; CAA76347.1;  
 DR EMBL, AL035413; CAB72321.1;  
 DR MIM, 603418;  
 DR InterPro: IPR001395; Aldo\_ket\_red.  
 DR Pfam: PF00248; Aldo\_ket\_red; 1.  
 KM Oxidoreductase.  
 FT ACT SITE 112 112 HYDROGEN-BOND DONOR (PROBABLE).  
 FT CONFLICT 113 113 A -> T (IN REF. 1).  
 SQ SEQUENCE 330 AA; 36618 MW; 3BBFB7ED0CAF4D54 CRC64;

Query Match 93.3%; Score 28; DB 1; Length 330;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 218 PVGRF 222

## RESULT 7

AR73\_HUMAN

ID AR73\_HUMAN STANDARD; PRT; 331 AA.  
AC 095154; 09NUC3;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE Aflatoxin B1 aldehyde reductase 2 (EC 1.-.-.-) (AFB1-AR 2).

GN AKR7A3.  
OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

OX NCBI\_TaxID=9606;  
[1]

RP SEQUENCE FROM N.A., AND CHARACTERIZATION.

RC TISSUE=Liver;

RA MEDLINE=99315412; PubMed=10383892;

RT "CDNA cloning, expression and activity of a second human aflatoxin B1-metabolizing member of the aldo-keto reductase superfamily,"

RT AKR7A3.",  
Carcinogenesis 20:1215-1223(1999).

RL Carcinogenesis 20:1215-1223(1999).

RN SEQUENCE FROM N.A.  
Hell R.;

CC Submitted (FE8-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING

CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIALCOHOL. COULD BE

CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC

CC EFFECTS OF AFLATOXIN B1.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; AF040639; AAD02195.1; -

CC EMBL; AL035413; CAB72322.1; -

CC InterPro: IPR001395; Aldo\_ket\_red.1.

CC Pfam: PF00248; aldo\_ket\_red.1.

CC PRINTS: PR00069; ALDKETREDTASE.

CC OXidoreductase.

CC ACT\_SITE 113 HYDROGEN-BOND DONOR (PROBABLE).

CC FT 113 E -> D (IN REF. 1).

CC FT 51 V -> M (IN REF. 1).

CC FT 138 A -> ADQSPGCGSFGWGLGPGADCCFPS (IN REF.

CC FT 201 2).

CC FT 215 N -> D (IN REF. 1).

CC FT 323 T -> A (IN REF. 1).

CC FT 331 AA; 37206 MW; B9C32C3C7102AB3 CRC64;

CC SEQUENCE

Query Match 93.3%; Score 28; DB 1; Length 331;  
Best Local Similarity 100.0%; Pred. No. 25;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||

DB 219 PVGRF 223

## RESULT 8

GIGL\_RHIME

ID GIGL\_RHIME STANDARD; PRT; 480 AA.  
AC P58393;

DT 01-MAR-2002 (Rel. 41, Created)

DT 01-MAR-2002 (Rel. 41, Last sequence update)

DE Glycogen synthase 1 (EC 2.4.1.21) (Starch [bacterial glycogen])

DE synthase 1)

GN GIGL OR R02846 OR SMC03924.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OX NCBI\_TaxID=382;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=1021;

RA MEDLINE=21396507; PubMed=11481430;

RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,

RA Bolstead P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,

RA Godrie T., Goffeau A., Kahn D., Kiss E., Leleure V., Masuy D.,

RA Renard C., Thebaud P., Puhler A., Purnelle B., Ramsperger U.,

RT "Analysis of the chromosome sequence of the legume symbiont

RT Sinorhizobium meliloti strain 1021."

RT Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).

CC -1- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.

CC -1- CATALYTIC ACTIVITY: ADP-glucose + ((1,4)-alpha-D-glucosyl)(N) -

CC ADP + ((1,4)-alpha-D-glucosyl)(N+1).

CC -1- PATHWAY: Glycogen biosynthesis; second step.

CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE

CC FAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
CC EMBL; AL591792; CAC47425.1; -

CC GIGL; AL591792; CAC47425.1; -

CC Complete proteome.

CC FT BINDING 15 ADP-GLUCOSE (BY SIMILARITY).

CC FT SEQUENCE 480 AA; 51408 MW; 84C584FE0564097 CRC64;

CC

Query Match 93.3%; Score 28; DB 1; Length 480;  
Best Local Similarity 100.0%; Pred. No. 37;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
|||||  
DB 56 PVGRF 60

RESULT 9

CATL\_BRAE

ID CATL\_BRAE STANDARD; PRT; 526 AA.

AC 09PT92; 0918V5;

DT 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DE 01-MAR-2002 (Rel. 41, Last annotation update)

DE Catalase (EC 1.11.1.6).

GN CAT.

OS Brachydanio rerio (zebrafish) (zebra danio).

OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;

OC Cypriniformes; Cyprinidae; Danio.

OK NCBI\_TaxID=7955;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RA Ken C.F., Lin C.T., Wu J.L., Shaw J.F.;  
 RT "Molecular cloning of a cDNA coding for catalase from zebrafish (Danio rerio).";  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RA Gerhard G.S., Kaufman E.J., Grundy M.A.;  
 RT "Molecular cloning and sequence analysis of the Danio rerio catalase gene.";  
 RL Comp. Biochem. Physiol. 127:447-457(2000).  
 CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC -1- COFACTOR: HEME GROUP AND NADP (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Peroxisomal (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AJ007505; CAB64949.1; -  
 DR EMBL; AF170069; AAF9686.1; -  
 DR HSSP; P00432; 4BLC.  
 DR ZFIN; ZDB-GENE-000210-20; cat.  
 DR InterPro; IPR002226; Catalase.  
 DR Pfam; PF00199; catalase.1.  
 DR PRINTS; PR00067; CATALASE.  
 DR PRODOM; PD000510; Catalase.1.  
 DR PROSITE; PS00437; CATALASE\_1; 1.  
 DR PROSITE; PS00438; CATALASE\_2; 1.  
 KW Oxidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KW Peroxisome; NADP.  
 FT ACT\_SITE 75 BY SIMILARITY.  
 FT ACT\_SITE 148 BY SIMILARITY.  
 FT BINDING 358 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 FT CONFLICT 110 V -> A (IN REF. 2).  
 FT CONFLICT 123 P -> S (IN REF. 2).  
 FT CONFLICT 152 T -> I (IN REF. 2).  
 FT CONFLICT 161 S -> F (IN REF. 2).  
 FT CONFLICT 350 MLQ -> NAA (IN REF. 2).  
 FT CONFLICT 478 M -> T (IN REF. 2).  
 SQ SEQUENCE 526 AA; 59654 MW; E1120D3796522785 CRC64;  
 OY 1 PVGRF 5  
 DB 312 PVGRF 316  
 Query Match 93.3%; Score 28; DB 1; Length 526;  
 Best Local Similarity 100.0%; Pred. NO. 41;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OK NCBI\_TaxID=562, 83334;  
 RN (1)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=95020533; PubMed=7934824;  
 RA Loubens I., Debarbieux L., Bohin A., Lacroix J.-M., Bohin J.-P.;  
 RT "Homology between a genetic locus (indA) involved in the osmoregulated biosynthesis of periplasmic glucans in Escherichia coli and a genetic locus (hrpM) controlling pathogenicity of Pseudomonas syringae.";  
 RL Mol. Microbiol. 10:329-340(1993).  
 RN (2)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 RN (3)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizobuchi K., Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N., Samped G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome corresponding to the 12.7-28.0 mln region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 RN (4)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Fofal G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobleck E.J., Davis N.W., Lin A., Dimalanta E.T., Potamousis K., Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C., Welch R.A., Blatter F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN (5)  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-O157:H7 / RIMD 0509952;  
 RX MEDLINE=21156231; PubMed=11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K., Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T., Iida T., Takami H., Honda T., Sasaki K., Ogasawara N., Yasunaga T., Kihara S., Shiba T., Hattori M., Shirogawa H.;  
 RT "Complete genome sequence of enterohaemorrhagic Escherichia coli O157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- FUNCTION: NECESSARY FOR NORMAL GLUCOSYLTRANSFERASE ACTIVITY. IT COULD BE A GLUCOSYL TRANSFERASE OR SIMPLY A SUBUNIT OF THIS ENZYME.  
 CC -1- PATHWAY: OPG (OSMOREGULATED PERIPLASMIC GLUCANS) BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -1- SIMILARITY: BELONGS TO THE GLUCOSYLTRANSFERASE FAMILY 2. STRONG, TO P. SYRINGAE HRP.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X64197; CAA5522.1; -

DR EMBL: AE000206; AAC74133.1; -

DR EMBL: D90742; BAA5848.1; -

DR EMBL: AE005315; AAG55795.1; -

DR EMBL: AF002555; BAB34850.1; ALT\_INIT.

DR PIR: S35418; S35418.

DR EcoGene: Egl1886; mdoH.

DR InterPro: IPR001173; Glycosyltransf\_2.

DR Pfam: PF00535; Glycosyltransf\_2; 1.

DR Transmembrane; Inner membrane; Transferase; Glycosyltransferase;

KM Complete proteome.

FT DOMAIN 1 139 CYTOPLASMIC (POTENTIAL).

FT TRANSMEM 140 160 POTENTIAL.

FT DOMAIN 161 193 CYTOPLASMIC.

FT TRANSMEM 194 214 POTENTIAL.

FT DOMAIN 215 425 PERIPLASMIC.

FT TRANSMEM 426 446 PERIPLASMIC.

FT DOMAIN 447 512 POTENTIAL.

FT TRANSMEM 513 533 CYTOPLASMIC.

FT DOMAIN 534 569 PERIPLASMIC.

FT TRANSMEM 570 590 POTENTIAL.

FT DOMAIN 591 614 CYTOPLASMIC.

FT TRANSMEM 615 635 POTENTIAL.

FT DOMAIN 636 679 PERIPLASMIC.

FT TRANSMEM 680 700 POTENTIAL.

FT DOMAIN 701 745 CYTOPLASMIC.

FT TRANSMEM 746 766 POTENTIAL.

FT DOMAIN 767 847 PERIPLASMIC.

FT CONFLICT 19 19 A -> G (IN REF. 1).

FT CONFLICT 289 289 P -> L (IN REF. 1).

SQ SEQUENCE 847 AA; 96937 MW; 7DCAF93640180944 CRC64;

Query Match 93.3%; Score 28; DB 1; Length 847;

Best Local Similarity 100.0%; Pred. No. 66;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5

Db 98 PVGRF 102

RESULT 11

ANPA\_MOUSE STANDARD; PRT; 1057 AA.

ID ANPA\_MOUSE 18293;

AC 01-NOV-1990 (Rel. 16, Created)

DT 01-FEB-1995 (Rel. 31, Last sequence update)

DE 16-OCT-2001 (Rel. 40, Last annotation update)

DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)

DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide

DE A-type receptor).

GN NPRL OR NPRA.

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J;

RX MEDLINE=90324219; PubMed=1973687;

RA Pandey K.N., Singh S.;

RT "Molecular cloning and expression of murine guanylate cyclase/atrial

RT natriuretic factor receptor cDNA.";

RL J. Biol. Chem. 265:12342-12348(1990).

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=95139994; PubMed=7838126;

RA Schoenfeld J.R., Sehl P., Quan C., Burnier J.P., Lowe D.G.;

RT "Agonist selectivity for three species of natriuretic peptide

RT receptor-A.";

RL Mol. Pharmacol. 47:172-180(1995).

CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE

CC CYCLASE ACTIVITY ON BINDING OF ANP.

CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS. TWO

CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)

CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE

CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.

CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE

CC FAMILY.

CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: J05504; AAA37670.1; -

DR PIR: I31932; AAA66945.1; -

DR PIR: A36568; OYMSAR.

DR PIR: A35088; A35088.

DR HSSP: Q02846; IAWL.

DR MGD: MGI:97371; NP1.

DR InterPro: IPR001170; ANF\_receptor.

DR InterPro: IPR001828; ANF\_receptor.

DR InterPro: IPR000719; Euk\_kinase.

DR InterPro: IPR001054; Guanylyl\_cyclase.

DR Pfam: PF01094; ANF\_receptor; 1.

DR Pfam: PF00211; guanylate\_cyc; 1.

DR Pfam: PF00069; kinase; 1.

DR PRINTS: PR00255; NATPEPTIDER.

DR SMART: SM00044; CYCC; 1.

DR PROSITE: PS00458; ANF\_RECEPTORS; 1.

DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.

DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.

DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.

KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;

KW GMP synthesis; Signal.

FT SIGNAL 1 28

FT CHAIN 29 1057

FT DOMAIN 29 469

FT TRANSMEM 470 490

FT DOMAIN 491 1057

FT DOMAIN 524 801

FT DOMAIN 872 1002

FT DISULFID 88 114

FT DISULFID 192 241

FT DISULFID 451 451

FT DISULFID 460 460

FT CARBOHYD 41 41

FT CARBOHYD 208 208

FT CARBOHYD 334 334

FT CARBOHYD 375 375

FT CARBOHYD 382 382

FT CARBOHYD 423 423

FT CONFLICT 3 3

FT CONFLICT 39 39

FT CONFLICT 122 122

FT CONFLICT 130 130

FT CONFLICT 285 285

FT CONFLICT 301 301

FT CONFLICT 404 405

FT CONFLICT 590 590

FT CONFLICT 652 652

FT CONFLICT 833 833

FT CONFLICT 958 958

FT CONFLICT 1044 1044

FT CONFLICT 1050 1050

ATRIAL NATRIURETIC PEPTIDE RECEPTOR A.

EXTRACELLULAR (POTENTIAL).

POTENTIAL.

CYTOPLASMIC (POTENTIAL).

PROTEIN KINASE-LIKE.

GUANYLATE CYCLASE.

BY SIMILARITY.

INTERCHAIN (PROBABLE).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

N-LINKED (GLCNAC. . .) (POTENTIAL).

G -> R (IN REF. 1).

L -> V (IN REF. 1).

G -> D (IN REF. 1).

V -> E (IN REF. 1).

R -> L (IN REF. 1).

A -> R (IN REF. 1).

FS -> SP (IN REF. 1).

H -> Q (IN REF. 1).

G -> C (IN REF. 1).

A -> P (IN REF. 1).

R -> S (IN REF. 1).

T -> S (IN REF. 1).

E -> D (IN REF. 1).

FT CONFLICT 1055 1057 TRG -> SRA (IN REF. 1).  
SQ SEQUENCE 1057 AA: 119109 MW: 53A54AFB2C8EF253 CRC64:

Query Match 93.3% Score 28; DB 1; Length 1057;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
11111  
Db 120 PVGRF 124

## RESULT 12

ANPA\_RAT STANDARD; PRT: 1057 AA.  
ID ANPA\_RAT  
AC P18910;  
DT 01-NOV-1990 (Rel. 16, Created)  
DT 01-NOV-1990 (Rel. 16, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)  
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide  
A-type receptor).  
DE A-type receptor).  
GN NPRL.  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RA MEDLINE=89143770; PubMed=2563900;  
RA Chinkers M., Garbers D.L., Chang M.S., Lowe D.G., Chn H.,  
RA Goeddel D.V., Schulz S.;  
RT "A membrane form of guanylate cyclase is an atrial natriuretic  
peptide receptor.";  
RL Nature 338:78-83(1989).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=91056089; PubMed=1978722;  
RA Yamaguchi M., Rutledge L.J., Garbers D.L.;  
RT "The primary structure of the rat guanylyl cyclase A/atrial  
natriuretic peptide receptor gene.";  
RL J. Biol. Chem. 265:20414-20420(1990).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=91352095; PubMed=1679239;  
RA Duda T., Goraczniak R.M., Shatta R.K.;  
RT "Site-directed mutational analysis of a membrane guanylate cyclase  
RT cDNA reveals the atrial natriuretic factor signaling site.";  
RL Proc. Natl. Acad. Sci. U.S.A. 88:7882-7886(1991).  
CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE  
CYCLASE ACTIVITY ON BINDING OF ANP.  
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO  
WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)  
WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE  
CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC  
DOMAIN OF PROTEIN KINASES.  
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
FAMILY.  
CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
-----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----

DR EMBL: X14773; CAA32881.1; -;  
DR EMBL: J05677; AAA41200.1; -;  
DR EMBL: M74535; AAA41202.1; -;  
DR PIR: S03348; OYRTR.  
DR HSSP: Q02846; 1AWL.  
DR InterPro: IPR001170; ANF\_receptor.  
DR InterPro: IPR001828; ANF\_receptor.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR001054; Guanylt\_cyclase.  
DR Pfam: PF01094; ANF\_receptor; 1.  
DR Pfam: PF00211; guanylate\_cyc; 1.  
DR Pfam: PF00069; pkinase; 1.  
DR PRINTS: PR00255; NATPEPTIDER.  
DR SMART: SM00044; CYCC; 1.  
DR PROSITE: PS00458; ANF\_RECEPTORS; 1.  
DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
KW CGMP synthetase; Signal.  
FT SIGNAL 1 28  
FT CHAIN 29 1057  
FT DOMAIN 29 469  
FT TRANSMEM 470 490  
FT DOMAIN 491 1057  
FT DOMAIN 524 801  
FT DOMAIN 872 1002  
FT DISULFID 88 114  
FT DISULFID 192 241  
FT DISULFID 451 451  
FT DISULFID 460 460  
FT CARBOHYD 41 41  
FT CARBOHYD 208 208  
FT CARBOHYD 304 334  
FT CARBOHYD 375 375  
FT CARBOHYD 382 382  
FT CARBOHYD 423 423  
FT CONFLICT 366 366  
FT CONFLICT 392 392  
SQ SEQUENCE 1057 AA: 118951 MW: 9EA9A8E85AC05816 CRC64;

Query Match 93.3% Score 28; DB 1; Length 1057;  
Best Local Similarity 100.0%; Pred. No. 83;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
11111  
Db 120 PVGRF 124

RESULT 13  
ANPA\_HUMAN STANDARD; PRT: 1061 AA.  
ID ANPA\_HUMAN  
AC P16066;  
DT 01-APR-1990 (Rel. 14, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)  
DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide  
A-type receptor).  
GN NPRL OR ANPRA.  
GN Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Kidney;  
RA MEDLINE=89356605; PubMed=2569967;  
RA Lowe D.G., Chang M.S., Hellmiss R., Chen E., Singh S., Garbers D.L.,  
RA Goeddel D.V.;  
RT "Human atrial natriuretic peptide receptor defines a new paradigm for

RT second messenger signal transduction.";  
 RL EMO J. 8:1377-1384(1989).  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Blood;  
 RX MEDLINE-98289600; PubMed-9618281;  
 RA Takahashi T., Nakayama T., Soma M., Isumi Y., Kamatsuse K.;  
 RT "Organization of the human natriuretic peptide receptor A gene.";  
 RL Biochem. Biophys. Res. Commun. 246:736-739(1998).  
 RN SEQUENCE FROM N.A.  
 RP Maeda N., Knowles J.W.;  
 RT "Identification of functional polymorphisms in noncoding regions of  
 RL the human natriuretic peptide receptor A gene.";  
 RL Submitted (SEP-1999) to the EMBL/GenBank/DBJ databases.  
 RN SEQUENCE OF 634-1048 FROM N.A.  
 RC TISSUE-Retina;  
 RX MEDLINE-95042574; PubMed-7954658;  
 RA Pardasarakhi K., Kutty R.K., Gentleman S., Krishna G.;  
 RT "Expression of mRNA for atrial natriuretic peptide receptor guanylate  
 RL cyclase (ANPRA) in human retina.";  
 RL Cell. Mol. Neurobiol. 14:1-7(1994).  
 CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE  
 CC CYCLASE ACTIVITY ON BINDING OF ANP.  
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO  
 CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)  
 CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE  
 CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
 CC FAMILY.  
 CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X15357; CNA33417.1; -  
 DR EMBL: AB010491; BAA31199.1; -  
 DR EMBL: AB010470; BAA31199.1; JOINED.  
 DR EMBL: AB010471; BAA31199.1; JOINED.  
 DR EMBL: AB010472; BAA31199.1; JOINED.  
 DR EMBL: AB010473; BAA31199.1; JOINED.  
 DR EMBL: AB010474; BAA31199.1; JOINED.  
 DR EMBL: AB010475; BAA31199.1; JOINED.  
 DR EMBL: AB010476; BAA31199.1; JOINED.  
 DR EMBL: AB010477; BAA31199.1; JOINED.  
 DR EMBL: AB010478; BAA31199.1; JOINED.  
 DR EMBL: AB010479; BAA31199.1; JOINED.  
 DR EMBL: AB010480; BAA31199.1; JOINED.  
 DR EMBL: AB010481; BAA31199.1; JOINED.  
 DR EMBL: AB010482; BAA31199.1; JOINED.  
 DR EMBL: AB010483; BAA31199.1; JOINED.  
 DR EMBL: AB010484; BAA31199.1; JOINED.  
 DR EMBL: AB010485; BAA31199.1; JOINED.  
 DR EMBL: AB010486; BAA31199.1; JOINED.  
 DR EMBL: AB010487; BAA31199.1; JOINED.  
 DR EMBL: AB010488; BAA31199.1; JOINED.  
 DR EMBL: AB010489; BAA31199.1; JOINED.  
 DR EMBL: AB010490; BAA31199.1; JOINED.  
 DR EMBL: AF190631; AAF01340.1; -  
 DR EMBL: S72628; AAD14112.1; -  
 DR F1R: S04459; OTHUAR.  
 DR HSP: 002846; IAWL.  
 DR MIM: 108960; -  
 DR InterPro: IPR001170; ANP\_receptor.  
 DR InterPro: IPR001828; ANP\_receptor.

DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001054; Guanylyl\_cyclase.  
 DR Pfam: PF01094; ANP\_receptor.1.  
 DR Pfam: PF00211; guanylate\_cyc.1.  
 DR Pfam: PF00069; pkinase.1.  
 DR PRINTS: PR00255; NATPEPTIDER.  
 DR SMART: SM00044; CYC.1.  
 DR PROSITE: PS00458; ANP-RECEPTORS.1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS00125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM.1.  
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
 KW GMP synthetis; Signal.  
 FT SIGNAL 1 32  
 FT CHAIN 33 1061  
 FT DOMAIN 33 473  
 FT TRANSMEM 474 494  
 FT DOMAIN 495 1061  
 FT DOMAIN 528 805  
 FT DOMAIN 876 1006  
 FT DISULFID 92 118  
 FT DISULFID 196 245  
 FT DISULFID 455 455  
 FT DISULFID 464 464  
 FT CARBOHYD 34 34  
 FT CARBOHYD 43 45  
 FT CARBOHYD 212 415  
 FT CARBOHYD 338 338  
 FT CARBOHYD 379 379  
 FT CARBOHYD 386 386  
 FT CARBOHYD 427 427  
 SQ SEQUENCE 1061 AA: 118918 MW: 118918 MW: E6B5BD0FCA32EF70D CRC64;  
 Query Match 93.3%; Score 28; DB 1; Length 1061;  
 Best local Similarity 100.0%; Pred. No. 83;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 PVGRF 5  
 Db 124 PVGRF 128  
 RESULT 14  
 ZAN\_PIG STANDARD; PRT; 2476 AA.  
 AC Q28983;  
 DT 01-NOV-1997 (Rel. 35, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Zonadhesin precursor.  
 GN ZAN.  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 823-830; 859-872; 883-890;  
 RP 920-925; 960-967; 1235-1244; 1349-1354; 1518-1532; 1624-1656;  
 RP 1658-1667; 1777-1795 AND 1914-1921.  
 RC STRAIN-MEISHAN; TISSUE-Testis;  
 RX MEDLINE-96064658; PubMed-7592795;  
 RA Hardy D.M., Garbers D.L.;  
 RT "A sperm membrane protein that binds in a species-specific manner to  
 RT the egg extracellular matrix is homologous to von Willebrand  
 RT factor.";  
 RL J. Biol. Chem. 270:26025-26028(1995).  
 CC -1- FUNCTION: BINDS IN A SPECIES-SPECIFIC MANNER TO THE ZONA PELLUCIDA  
 CC OF THE EGG. MAY BE INVOLVED IN GAMETE RECOGNITION AND/OR  
 CC SIGNALING.  
 CC -1- SUBUNIT: PROBABLY FORMS COVALENT OLIGOMERS.  
 CC -1- SUBCELLULAR LOCATION: TYPE I MEMBRANE PROTEIN, EXCLUSIVELY ON THE  
 CC APICAL REGION OF THE SPERM HEAD (BY SIMILARITY).



CC	-1-	TISSUE SPECIFICITY: IN TESTIS, PRIMARILY IN HAPLOID SPERMATIDS.
CC	-1-	NOT IN LONG LAYER, HEART, SPLEEN, BRAIN, KIDNEY, EPIDIDYMUS.
CC	-1-	DOMAIN: THE MAM DOMAINS PROBABLY MEDIATE SPERM ADHESION TO THE ZONA PELLUCIDA.
CC	-1-	DOMAIN: DURING SPERM MIGRATION THROUGH THE REPRODUCTIVE TRACTS, THE MUCIN-LIKE DOMAIN MIGHT INHIBIT INAPPROPRIATE TRAPPING OF SERMATOCOA OR PROMOTING ADHESION TO THE OVICYTAL ISTHMUS.
CC	-1-	OLIGOMERIZATION (BY SIMILARITY TO HUMAN INTESTINAL MUCIN MUCC2).
CC	-1-	PWM: THE MAM DOMAIN AND THE MUCIN-LIKE DOMAINS ARE MISSING FROM THE ZONADHESIN THAT BINDS TO THE EGG EXTRACELLULAR MATRIX.
CC		PROCESSING MIGHT OCCUR DURING SPERM MATURATION AND/OR CAPACITATION.
CC	-1-	SIMILARITY: CONTAINS 2 MAM DOMAINS.
CC	-1-	SIMILARITY: CONTAINS 4.5 VWFD DOMAINS.
CC	-1-	SIMILARITY: CONTAINS 1 EGF-LIKE DOMAIN.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC		-----
DR	EMBL: U40024; AAC48486.1; -	
DR	InterPro: IPR000561; EGF-like.	
DR	InterPro: IPR000998; MAM.	
DR	InterPro: IPR002919; TIL.	
DR	InterPro: IPR003328; Tila.	
DR	InterPro: IPR001007; VWFC.	
DR	InterPro: IPR001846; Vwd.	
DR	pfam: PF00629; MAM; 2.	
DR	pfam: PF01826; TIL; 5.	
DR	pfam: PF02345; Tila; 5.	
DR	pfam: PF00094; vwd; 4.	
DR	SMART: SM00181; EGF; 1.	
DR	SMART: SM00137; MAM; 1.	
DR	SMART: SM00214; WVC; 2.	
DR	SMART: SM00216; VND; 4.	
DR	PROSITE: PS00022; EGF_1; 1.	
DR	PROSITE: PS01186; EGF_2; 4.	
DR	PROSITE: PS00740; MAM_1; 1.	
DR	PROSITE: PSS0060; MAM_2; 2.	
KW	Signal; Glycoprotein; Transmembrane; EGF-like domain; Cell adhesion; Repeat.	
FT	SIGNAL	1 29 POTENTIAL.
FT	CHAIN	30 2476 ZONADHESIN.
FT	DOMAIN	30 2418 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM	2419 2439 POTENTIAL.
FT	DOMAIN	2440 2476 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN	31 144 MAM 1.
FT	DOMAIN	147 312 MAM 2.
FT	DOMAIN	319 687 53 x HEPTAPEPTIDE REPEATS (APPROXIMATE) (MUCIN-LIKE DOMAIN).
FT	DOMAIN	688 799 VWFD 1 (PARTIAL).
FT	DOMAIN	800 1184 VWFD 2.
FT	DOMAIN	1185 1573 VWFD 3.
FT	DOMAIN	1574 1968 VWFD 4.
FT	DOMAIN	1969 2370 VWFD 5.
FT	DOMAIN	2366 2402 EGF-LIKE.
FT	DISULFID	2370 2381 BY SIMILARITY.
FT	DISULFID	2375 2390 BY SIMILARITY.
FT	DISULFID	2392 2401 BY SIMILARITY.
FT	CARBOHYD	109 109 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	269 269 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	735 735 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	758 758 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	833 833 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	1154 1154 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	1329 1329 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	1448 1448 N-LINKED (GLCNAC . . ) (POTENTIAL).
FT	CARBOHYD	1544 1544 N-LINKED (GLCNAC . . ) (POTENTIAL).

FT	CARBOHYD	1596	1596	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	1654	1654	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	1843	1843	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	1965	1965	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	2122	2122	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	2165	2165	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	2178	2178	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	2329	2329	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CARBOHYD	2359	2359	N-LINKED	(GLCNAC . . .)	(POTENTIAL)
FT	CONFLICT	823	823	C -> V (IN REF. 1; AA SEQUENCE).		
FT	CONFLICT	923	923	S -> Y (IN REF. 1; AA SEQUENCE).		
FT	CONFLICT	965	965	W -> K (IN REF. 1; AA SEQUENCE).		
FT	CONFLICT	1241	1241	S -> K (IN REF. 1; AA SEQUENCE).		
SQ	SEQUENCE	2476	270364	AA; MW; A13B690375A6548C CRC64;		

```

Query Match          93.3%; Score 28; DB 1; Length 2476;
Best Local Similarity 100.0%; Pred. NO. 2e+02;
Matches      5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY      1 PVGRF 5
        |||||
Db      1918 PVGRF 1922

RESULT 15
ID      YP15_STAAU
AC      YP15_STAAU STANDARD; PRT; 140 AA.
DT      01-JAN-1990 (Rel. 13, Created)
DT      01-JAN-1990 (Rel. 13, Last sequence update)
DT      01-FEB-1995 (Rel. 31, Last annotation update)
DE      Hypothetical 15.5 kDa protein.
OS      Staphylococcus aureus.
OG      Staphid PSK1.
OC      Bacteria; Firmicutes; Bacillus/Clostridium group;
OC      Bacillus/Staphylococcus group; Staphylococcus.
OX      NCBI_TaxID=1280;
RN      [1]
RP      SEQUENCE FROM N.A.
RX      MEDLINE=89343620; PubMed=2548057;
RA      Rouch D.A., Messerotti L.J., Loo L.S.L., Jackson C.A., Skurray R.A.;
RT      "Trimethoprim resistance transposon Tn4003 from Staphylococcus aureus
RT      encodes genes for a dilydrofolate reductase and thymidylate
RT      synthetase flanked by three copies of IS257."
RL      Mol. Microbiol. 3:161-175(1989).
CC      -1 SIMILARITY: TO B.SUBTILIS DECV.
CC      -----
CC      This SWISS-PROT entry is copyright. It is produced through a collaboration
CC      between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC      the European Bioinformatics Institute. There are no restrictions on ways
CC      CC use by non-profit institutions as long as its content is in no way
CC      modified and this statement is not removed. Usage by and for commercial
CC      entities requires a license agreement (see http://www.isb-sb.ch/announce/
CC      or send an email to license@isb-sib.ch).
CC      -----
CC      EMBL; X13290; CAA31650.1; -.
CC      PIR; S04165; S04165.
CC      Hypothetical protein; Plasmid.
KM      SEQUENCE 140 AA; 15462 MW; 3FCC0CB110239C6F CRC64;

```

```

Query Match          90.0%; Score 27; DB 1; Length 140;
Best Local Similarity 80.0%; Pred. No. 18;
Matches 4; Conservative 1; Mismatches 0; Indels 0; Gaps 0.

QY      1 PVGRF 5
        | : |||
        65 PIGRF 69


```

Job time: 1134 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:17 ; Search time 311.85 Seconds  
(without alignments)  
3.883 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_16\_22  
Perfect score: 30  
Sequence: 1 PVGRFX 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	28	93.3	106	2	067984
2	28	93.3	118	4	09P195
3	28	93.3	141	3	096W56
4	28	93.3	154	2	054139
5	28	93.3	161	11	092221
6	28	93.3	163	4	09H7F5
7	28	93.3	170	16	055547
8	28	93.3	173	16	P74584
9	28	93.3	193	12	09OJ59
10	28	93.3	198	5	09VNS8
11	28	93.3	200	4	09NPV7
12	28	93.3	257	16	09HW71
13	28	93.3	257	16	09BNM6
14	28	93.3	258	11	0921C1
15	28	93.3	260	2	0880C8
16	28	93.3	294	2	006001

17	28	93.3	294	10	092085	09u85 arabidopsis
18	28	93.3	298	2	034287	034287 zymomonas m
19	28	93.3	299	16	P95060	P95060 mycobacteri
20	28	93.3	313	5	09UBB2	09UBB2 hexamita sp
21	28	93.3	315	11	09DCM1	09DCM1 mus musculu
22	28	93.3	357	5	09NGT9	09NGT9 leishmania
23	28	93.3	357	5	09NE10	09NE10 leishmania
24	28	93.3	358	2	09L1H6	09L1H6 streptomyce
25	28	93.3	366	11	09DI57	09DI57 mus musculu
26	28	93.3	367	2	087693	087693 bacillus me
27	28	93.3	383	10	09SP62	09SP62 impatiens b
28	28	93.3	420	2	093E77	093E77 streptomyce
29	28	93.3	424	17	029070	029070 archaeoglob
30	28	93.3	445	2	050531	050531 streptomyce
31	28	93.3	458	4	09BV03	09BV03 homo sapien
32	28	93.3	458	4	09NVR4	09NVR4 homo sapien
33	28	93.3	501	3	008229	008229 saccharomyc
34	28	93.3	511	3	059714	059714 schizosacch
35	28	93.3	526	13	0918V5	0918V5 brachydanio
36	28	93.3	559	16	09KBE5	09KBE5 bacillus ha
37	28	93.3	585	17	029566	029566 archaeoglob
38	28	93.3	601	17	030039	030039 archaeoglob
39	28	93.3	612	2	09X8C4	09X8C4 streptomyce
40	28	93.3	647	4	09BWM3	09BWM3 homo sapien
41	28	93.3	648	4	09P209	09P209 homo sapien
42	28	93.3	669	4	09BSK4	09BSK4 homo sapien
43	28	93.3	931	3	09UJZ3	09UJZ3 aspergillus
44	28	93.3	937	3	074168	074168 aspergillus
45	28	93.3	1057	11	091X04	091X04 mus musculu

## ALIGNMENTS

RESULT 1  
ID 067984 PRELIMINARY: PRT: 106 AA.  
AC 067984;  
DT 01-AUG-1998 (TREMBLrel. 07, Created)  
DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE 3-OXOADIPYL COA THIOLASE HOMOLOG (EC 2.3.1.) (FRAGMENT).  
GN PCAF.  
OS Rhodococcus opacus (Nocardia opaca).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Nocardiaceae; Rhodococcus.  
OX NCBI\_TaxID=37919;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-1CP;  
RX MEDLINE=98155132; PubMed=9495744;  
RA Eulberg D., Lakner S., Golovleva L.A., Schlomann M.;  
RT "Characterization of a proteobacterial catabolic gene cluster from Rhodococcus opacus 1CP: evidence for a merged enzyme with 4-carboxymuconolactone-decarboxylating and 3-oxoadipate enol-lactone-hydrolyzing activity.";  
RL J. Bacteriol. 180:1072-1081(1998).  
DR EMBL: AF003947; AAC38248.1; -.  
DR HSSP: P27796; JAFY.  
DR InterPro: IPR002155; Thiolase.  
DR Pfam: PF00108; Thiolase; 1.  
KW Transferase; Acyltransferase.  
FT NON\_TER  
SQ SEQUENCE 106 AA; 10856 MW; 7C0A24EDE86E9C2F CMC64;

Query Match 93.3%; Score 28; DB 2; Length 106;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 18 PVGRF 22

RESULT 2  
ID 09P195 PRELIMINARY: PRT: 118 AA.  
AC 09P195: 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE 01-OCT-2000 (TREMBlrel. 15, Last annotation update)  
DE PRO1722.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=LIVER;  
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,  
He F.;  
RT "Functional prediction of the coding sequences of 79 new genes deduced  
by analysis of cDNA clones from human fetal liver."  
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF119851; AAF69605.1; -  
SQ SEQUENCE 118 AA; 13067 MW; D6716A242C773055 CRC64;

Query Match 93.3%; Score 28; DB 4; Length 118;  
Best Local Similarity 100.0%; Pred. No. 35;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 94 PVGRF 98

RESULT 3  
ID 096W56 PRELIMINARY: PRT: 141 AA.  
AC 096W56: 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE YIP3-LIKE PROTEIN (FRAGMENT).  
GN YIP3.  
OS Candida albicans (Yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; mitosporic Saccharomycetales; Candida.  
OX NCBI\_TaxID=5476;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=B207;  
RA Burke T.J., Rhoads D.D.;  
RT "Evolution of ribosomal protein S14 gene structure: Candida albicans,  
Schizosaccharomyces pombe, and selected ascomycetous fungi."  
RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF365406; AAK60139.1; -  
FT NON\_TER 1  
SQ SEQUENCE 141 AA; 15509 MW; 36F680DFD7D76419 CRC64;

Query Match 93.3%; Score 28; DB 3; Length 141;  
Best Local Similarity 100.0%; Pred. No. 43;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 81 PVGRF 85

RESULT 4  
ID 054139 PRELIMINARY: PRT: 154 AA.  
AC 054139;

DT 01-JUN-1998 (TREMBlrel. 06, Created)  
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
DE HYPOTHETICAL 16.9 KDA PROTEIN.  
GN SC2B9.14.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Oliver K., Harris D.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Parkhill J., Barrall B.G., Rajandream M.A.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,  
Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RL MOL. Microbiol. 21:77-96(1996).  
DR EMBL: AL021530; CA16482.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 154 AA; 16884 MW; 9D9B47F368B4CD03 CRC64;

Query Match 93.3%; Score 28; DB 2; Length 154;  
Best Local Similarity 100.0%; Pred. No. 47;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 32 PVGRF 36

RESULT 5  
ID 092221 PRELIMINARY: PRT: 161 AA.  
AC 092221: 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE SIMILAR TO HYPOTHETICAL PROTEIN FROM CLONE 24796.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Strausberg R.;  
RL Submitted (APR-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: BC006689; AAH06689.1; -  
SQ SEQUENCE 161 AA; 17643 MW; AB2527A56EBA6CF9 CRC64;

Query Match 93.3%; Score 28; DB 11; Length 161;  
Best Local Similarity 100.0%; Pred. No. 49;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
DB 40 PVGRF 44

RESULT 6  
Q9H7F5

ID 09H7E5 PRELIMINARY; PRT; 163 AA.  
AC 09H7E5;  
DT 01-MAR-2001 (TrEMBLrel. 16, Created)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)  
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)  
DE CDNA: FLJ20974 FIS. CLONE ADS001596.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=ADIPOSE TISSUE;  
RA Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T., Matsumura K.,  
RA Nakajima Y., Mizuno T., Morinaga M., Tanigami A., Fujiwara T., Ono T.,  
RA Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y., Ota T., Suzuki Y.,  
RA Obayashi M., Mishit T., Shibahara T., Tanaka T., Nakamura Y.,  
RA Isogai T., Sugano S.,  
RT MEDO human cDNA sequencing project."  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AK024627; BAI14937.1; -  
SO SEQUENCE 163 AA; 17712 MW; 7CDBD8306DE81EDC CRC64;

Query Match 93.3%; Score 28; DB 4; Length 163;  
Best Local Similarity 100.0%; Pred. No. 50;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
DB 43 PVGRF 47

RESULT 7  
ID 055547 PRELIMINARY; PRT; 170 AA.  
AC 055547;  
DT 01-NOV-1996 (TrEMBLrel. 01, Created)  
DT 01-NOV-1996 (TrEMBLrel. 01, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 18.6 KDA PROTEIN.  
GN STL0293.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=PCC 6803;  
RA Tabata S.;  
RL Submitted (Aug-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=96127529; PubMed=8590279;  
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugita M., Tabata S.;  
RT Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. I. Sequence features in the 1 Mb  
RT region from map positions 64k to 92k of the genome.";  
RL DNA Res. 2:153-166(1995).  
RN [3]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Natsu K., Okumura S.,  
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL; D63999; BAI10045.1; -  
KM Hypothetical protein; Complete proteome.

SQ SEQUENCE 170 AA; 18583 MW; 6BC7097A9BB3CBF3 CRC64;

Query Match 93.3%; Score 28; DB 16; Length 170;  
Best Local Similarity 100.0%; Pred. No. 52;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
DB 47 PVGRF 51

RESULT 8  
ID P74584 PRELIMINARY; PRT; 173 AA.  
AC P74584;  
DT 01-FEB-1997 (TrEMBLrel. 02, Created)  
DT 01-FEB-1997 (TrEMBLrel. 02, Last sequence update)  
DT 01-JUN-2000 (TrEMBLrel. 14, Last annotation update)  
DE HYPOTHETICAL 20.5 KDA PROTEIN.  
GN SLR0667.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
NCBI\_TaxID=1148;  
RN [1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Muraki A., Nakazaki N., Natsu K., Okumura S.,  
RA Hosouchi T., Matsuno A., Sugita M., Sasamoto S., Kimura T.,  
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL; D90916; BAI18691.1; -  
KM Hypothetical protein; Complete proteome.  
SO SEQUENCE 173 AA; 20509 MW; 2E2414F099C8B2F7 CRC64;

Query Match 93.3%; Score 28; DB 16; Length 173;  
Best Local Similarity 100.0%; Pred. No. 53;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
DB 95 PVGRF 99

RESULT 9  
ID 090J59 PRELIMINARY; PRT; 193 AA.  
AC 090J59;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE STRAIN 229, COMPLETE GENOME.  
GN B4.  
OS Human herpesvirus 6B.  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Roseolovirus.  
NCBI\_TaxID=32604;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=229;  
RX MEDLINE=99412318; PubMed=10482553;  
RA Dominguez G., Dambaugh T.R., Stamey F.R., Dewhurst S., Inoue N.,  
RA Pellett P.E.;  
RT "human herpesvirus 6B genome sequence: coding content and comparison  
RT with human herpesvirus 6A";  
RL J. Virol. 73:8040-8052(1999).  
DR EMBL; AF157706; AAD49620.1; -

SQL SEQUENCE 193 AA; 21504 MW; 728DF48DE420B7A4 CRC64;

Query Match 93.3%; Score 28; DB 12; Length 193;

Best Local Similarity 100.0%; Pred. No. 59;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
DB 106 PVGRF 110

RESULT 10

Q9VNS8 PRELIMINARY; PRT; 198 AA.

AC Q9VNS8; 01-MAY-2000 (TREMblrel. 13, Created)

DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)

DE 01-MAY-2000 (TREMblrel. 13, Last annotation update)

CG14456 PROTEIN.

OS CG14456.

OC Drosophila melanogaster (Fruit fly).

OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

OC Ephydroidea; Drosophilidae; Drosophila.

NCBI\_TaxID=7227;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN-BERKELEY;

RA MEDLINE-20196006; PubMed-10731132;

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,

RA Amaratunga P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,

RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,

RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe C.M., Pfeiffer B.D.,

Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,

Abdill J.F., Agbayani A., An H.-J., Andrews-Efankoch C., Baldwin D.,

Balieu R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,

Beeson K.Y., Benos P.V., Bertan B.P., Bhattacharya D., Bolshakov S.,

Borokova D., Botchan M.A., Bouck J., Brockett P., Broctier P.,

Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,

Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,

de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,

Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,

Dutrin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,

Foster C., Gabrielian A.E., Gary N.S., Gelbart W.M., Glasser K.,

Glock A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,

Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,

Hoslin D., Houston K.A., Howland T.J., Mei M.-H., Ibegwam C.,

Jatelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,

Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

Laake P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,

Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,

Mount S.M., Muhlhorn N.V., Mobarry C., Morris J., Moshrefi A.,

Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclio J.M.,

Palazzo M., Pittman G.S., Pan S., Pollard J., Pui V., Reese M.G.,

Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,

Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,

Spieler E., Spradling A.C., Stapleton M., Strong R., Sun E.,

Sylvester R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,

Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,

Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,

Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang Q., Zhao Q., Zheng L.,

Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,

Query Match 93.3%; Score 28; DB 5; Length 198;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
DB 104 PVGRF 108

RESULT 11

Q9NP7 PRELIMINARY; PRT; 200 AA.

AC Q9NP7; 01-OCT-2000 (TREMblrel. 15, Created)

DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE SIMILAR TO (NP\_034322.1) SEX-DETERMINATION PROTEIN HOMOLOG FEMA1

(FRAGMENT).

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

NCBI\_TaxID=9606;

RP [1]

RP SEQUENCE FROM N.A.

RA Carim L., Estivill X., Escarceller M., Sumoy L.,

Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.

RP [2]

RP SEQUENCE FROM N.A.

RA Auffray C., Ansoorge W., Ballabio A., Estivill X., Gibson K.,

Lehach H., Poustka A., Lundberg J.,

"The European IMAGE consortium for Integrated Molecular analysis of

human gene transcripts."

Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.

EMBL: AL365408; CAB96952.1;

InterPro: IPR02110; ANK.

DR SMART; SM00248; ANK; 2.

DR PROSITE; PS50088; ANK\_REPEAT; 2.

DR PROSITE; PS50297; ANK\_REPEAT\_REGION; 1.

FT ANK repeat; Repeat.

FT NON\_TER

SQL SEQUENCE 200 AA; 22361 MW; 0443BC897D7E1238 CRC64;

Query Match 93.3%; Score 28; DB 4; Length 200;

Best Local Similarity 100.0%; Pred. No. 62;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 PVGRF 5  
11111  
DB 85 PVGRF 89

RESULT 12

Q9HW71 PRELIMINARY; PRT; 257 AA.

AC Q9HW71; 01-MAR-2001 (TREMblrel. 16, Created)

DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)

DE 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE PROBABLE ENOYL-COA HYDRATASE/ISOMERASE.

GN PA4330.

OS Pseudomonas aeruginosa.

OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;

OC Pseudomonas.

NCBI\_TaxID=287;

RP [1]

RP SEQUENCE FROM N.A.

RC STRAIN-ATCC 15692 / PA01;

RA MEDLINE-20437337; PubMed-10984043;

RA Stover C.K., Pham X.-Q.T., Ervin A.L., Mizoguchi S.D., Warren P.,

Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,

Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,

RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an  
RT opportunistic pathogen."  
RL Nature 406:959-964(2000).  
DR EMBL: AE004849; AAC07718.1; -  
KW Isomerase; Complete proteome.  
SQ SEQUENCE 257 AA; 28152 MW; FACFCEBA49D1BB7 CRC64;

Query Match 93.3%; Score 28; DB 16; Length 257;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 80 PVGRF 84

RESULT 13  
O98NM6 PRELIMINARY; PRT; 257 AA.  
AC O98NM6:  
DT 01-OCT-2001 (TREMBlrel. 18, Created)  
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMBlrel. 18, Last annotation update)  
DE 2,4-DIHYDROXYHEPT-2-ENE-1,7-DIOIC ACID ALDOLASE.  
GN MRA0073.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idesawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
RT Mesorhizobium loti."  
RL DNA Res. 7:331-338(2000).  
DR EMBL: AP002994; BAB47735.1; -  
KW Complete proteome.  
SQ SEQUENCE 257 AA; 27314 MW; 4213E3C647342C06 CRC64;

Query Match 93.3%; Score 28; DB 16; Length 257;  
Best Local Similarity 100.0%; Pred. No. 80;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 70 PVGRF 74

RESULT 14  
O921C1 PRELIMINARY; PRT; 258 AA.  
AC O921C1:  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CONNEXIN29.  
GN CX29.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.

RX MEDLINE=21392673; PubMed=11501764;  
RA Soehl G., Elberger J., Jung Y., Kozak C., Willecke K.;  
RT "The mouse gap junction gene Connexin29 is highly expressed in  
RT scientific nerve and regulated during brain development."  
RL Biol. Chem. 382:973-978(2001).  
DR EMBL: AJ297318; CAC29245.1; -  
SQ SEQUENCE 258 AA; 28982 MW; 26D13AB3AC009458 CRC64;

Query Match 93.3%; Score 28; DB 11; Length 258;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 19 PVGRF 23

RESULT 15  
O88068 PRELIMINARY; PRT; 260 AA.  
AC O88068:  
DT 01-NOV-1998 (TREMBlrel. 08, Created)  
DT 01-NOV-1998 (TREMBlrel. 08, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE PUTATIVE DEHYDROGENASE.  
GN SC135.33C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Oliver K., Harris D.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RA Parkhill J., Barrell B.G., Rastam W.A.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
RL Mol. Microbiol. 21:77-96(1996).  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
DR EMBL: AL031541; GAA20822.1; -  
DR HSSP; P19992; HDC.  
DR InterPro; IPR002198; ADH\_short.  
DR Pfam; PF00106; adh\_short; 1.  
DR PRINTS; PRO0080; SDRFAMILY.  
KW Oxidoreductase.  
SQ SEQUENCE 260 AA; 27348 MW; B6A29754AE97026A CRC64;

Query Match 93.3%; Score 28; DB 2; Length 260;  
Best Local Similarity 100.0%; Pred. No. 81;  
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 PVGRF 5  
|||||  
DB 216 PVGRF 220

Search completed: September 13, 2002, 09:29:18  
Job time: 1063 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:34 ; Search time 399.68 Seconds  
(without alignments)  
5.558 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_2\_21  
Perfect score: 103  
Sequence: 1 PDINPAMYXXNGIRPVGRFX 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

1:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1980.DAT:*
2:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1981.DAT:*
3:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1982.DAT:*
4:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1983.DAT:*
5:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1984.DAT:*
6:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1985.DAT:*
7:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1986.DAT:*
8:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1987.DAT:*
9:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1988.DAT:*
10:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1989.DAT:*
11:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1990.DAT:*
12:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1991.DAT:*
13:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1992.DAT:*
14:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1993.DAT:*
15:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1994.DAT:*
16:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1995.DAT:*
17:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1996.DAT:*
18:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1997.DAT:*
19:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA1998.DAT:*
20:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2000.DAT:*
21:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2001.DAT:*
22:	/SIDSI/gcgdata/hold-geneeq/geneeqp-emb1/AA2001.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	97.1	20	AAW31394	Human type G prote
2	100	97.1	20	AAW97236	Human type G prote
3	100	97.1	20	AAW10365	Human type G prote
4	100	97.1	20	AAW49294	Human type G prote
5	100	97.1	20	AAW62534	Human type G prote
6	100	97.1	20	AAW90992	Human type G prote
7	100	97.1	21	AAW10366	Human type G prote
8	100	97.1	21	AAW10366	Human type G prote
9	100	97.1	21	AAW62535	Human type G prote
10	100	97.1	22	AAW31396	Human type G prote
11	100	97.1	22	AAW10367	Human type G prote

12	100	97.1	22	AAW62536	Human CRH releasin
13	100	97.1	31	AAW31391	Human type G prote
14	100	97.1	31	AAW97235	Human type G prote
15	100	97.1	31	AAW87615	Human 19P2 ligand
16	100	97.1	31	AAW10362	Human 19P2 ligand
17	100	97.1	31	AAW49291	Human 19P2 ligand
18	100	97.1	31	AAW62531	Human 19P2 ligand
19	100	97.1	31	AAW90991	Human CRH releasin
20	100	97.1	31	AAW90995	Human CRH releasin
21	100	97.1	32	AAW31392	Human type G prote
22	100	97.1	32	AAW10363	Human type G prote
23	100	97.1	32	AAW62532	Human CRH releasin
24	100	97.1	33	AAW31393	Human type G prote
25	100	97.1	33	AAW10364	Human type G prote
26	100	97.1	33	AAW62533	Human CRH releasin
27	100	97.1	87	AAW31390	Human type G prote
28	100	97.1	87	AAW97226	Human type G prote
29	100	97.1	87	AAW10361	Human type G prote
30	100	97.1	87	AAW62530	Human CRH releasin
31	99	96.1	20	AAW31387	Human CRH releasin
32	99	96.1	20	AAW31374	Human CRH releasin
33	99	96.1	20	AAW97232	Human CRH releasin
34	99	96.1	20	AAW97234	Human CRH releasin
35	99	96.1	20	AAW95191	Human CRH releasin
36	99	96.1	20	AAW95175	Human CRH releasin
37	99	96.1	20	AAW10350	Human CRH releasin
38	99	96.1	20	AAW10358	Human CRH releasin
39	99	96.1	20	AAW49301	Human CRH releasin
40	99	96.1	20	AAW49302	Human CRH releasin
41	99	96.1	20	AAW62519	Human CRH releasin
42	99	96.1	20	AAW62527	Human CRH releasin
43	99	96.1	20	AAW90994	Human CRH releasin
44	99	96.1	20	AAW90996	Human CRH releasin
45	99	96.1	20	AAW69554	Human CRH releasin

#### ALIGNMENTS

RESULT	1
AAW31394	standard; Peptide: 20 AA.
ID	AAW31394
XX	AAW31394;
AC	AAW31394;
XX	AAW31394;
DT	06-APR-1998 (first entry)
XX	06-APR-1998 (first entry)
DE	Human type G protein-coupled receptor ligand fragment 4.
XX	Human type G protein-coupled receptor ligand fragment 4.
KW	G protein-coupled receptor; ligand binding; pharmaceutical;
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;
KW	therapeutic agent.
XX	
OS	Homo sapiens.
XX	Homo sapiens.
PN	WO9724436-A2.
XX	WO9724436-A2.
PD	10-JUL-1997.
XX	10-JUL-1997.
PF	26-DEC-1996; 96WO-JP03821.
XX	26-DEC-1996; 96WO-JP03821.
PR	18-SEP-1996; 96JP-0246573.
PR	18-SEP-1996; 96JP-0246573.
PR	28-DEC-1995; 95JP-0343371.
PR	28-DEC-1995; 95JP-0343371.
PR	15-MAR-1996; 96JP-0059419.
PR	15-MAR-1996; 96JP-0059419.
XX	12-AUG-1996; 96JP-0211805.
XX	12-AUG-1996; 96JP-0211805.
PA	(TAKE ) TAKEDA CHEM IND LTD.
XX	(TAKE ) TAKEDA CHEM IND LTD.
PI	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,
PI	Kawamata Y, Kitada C;
XX	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,
XX	Kawamata Y, Kitada C;
DR	WPI: 1997-363672/33.
DR	N-PSDB; AAV02431.

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 185; 258pp; English.  
 XX

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
 CC sequence represented in AAW91390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polypthagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC Turner's syndrome, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligogalactia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX

SQ Sequence 20 AA:

Query Match 97.1%; Score 100; DB 18; Length 20;  
 Best Local Similarity 89.5%; Pred. No. 1.8e-10;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPYGRF 19  
 |||||  
 DB 2 pdlnpawyasrglrpygrf 20

RESULT 2

AAW97236 standard; peptide; 20 AA.

AAW97236;

06-MAY-1999 (first entry)

Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypocoarctanism; gonocyst cecogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenorrhoea; galactorrhea;  
 KW prolactinoma; infertility; amenorrhea; galactosporia;  
 KW acromegaly; Chlari-Frommel syndrome; Argon-z-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; chorioncarcinoma; hydralid mole;  
 KW interruption mole; abortion; unthriftly fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 XX

Homo sapiens.

WO9858962-A1.

30-DEC-1998.

22-JUN-1998; 98WO-JP02765.

23-JUN-1997; 97JP-0165437.

(TAKE ) TAKEDA CHEM IND LTD.

Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX

PS Claim 3; Page 166; 241pp; English.

XX The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypocoarctanism, gonocyst cecogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenorrhoea, galactorrhea,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactosporia,  
 CC acromegaly, Chlari-Frommel syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC chorioncarcinoma, hydralid mole, interruption mole, unthriftly fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 XX

SQ Sequence 20 AA:

Query Match 97.1%; Score 100; DB 20; Length 20;  
 Best Local Similarity 89.5%; Pred. No. 1.8e-10;  
 Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPYGRF 19  
 |||||  
 DB 2 pdlnpawyasrglrpygrf 20

RESULT 3

AAW10365 standard; peptide; 20 AA.

AAW10365;

24-NOV-2000 (first entry)

Human oxytocin secretion promoting peptide SEQ ID NO: 35.

XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactosporia; goat; pig;  
 KW veterinary medicine; milk production.  
 XX

Homo sapiens.

WO200038704-A1.

06-JUL-2000.

22-DEC-1999; 99WO-JP07199.

25-DEC-1998; 98JP-0369585.

(TAKE ) TAKEDA CHEM IND LTD.

Matsumoto H, Kitada C, Hinuma S;

WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine

PS Disclosure: Page 63; 72pp; Japanese.  
 CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.,  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.  
 CC  
 CC  
 SS Sequence 20 AA:  
 OY 1 PDINPAMYXXRGIRPVGRF 19  
 Db 2 pdinpamyasrgirpvgrf 20  
 OY 1 PDINPAMYXXRGIRPVGRF 19  
 Db 2 pdinpamyasrgirpvgrf 20  
 RESULT 4  
 AAY49294  
 ID AAY49294 standard; peptide: 20 AA.  
 XX  
 AC AAY49294:  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE 19P2 ligand peptide fragment.  
 XX  
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 20  
 FT /note="C-terminal amide"  
 XX  
 PN WO9960112-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 20-MAY-1999; 99WO-JP02650.  
 XX  
 PR 21-MAY-1998; 98JP-0140293.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Matsumoto H, Kikada C, Hinuma S;  
 XX  
 DR WPI: 2000-039381/03.  
 XX  
 PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 XX  
 PS Disclosure: Page 26; 73pp; Japanese.  
 CC  
 CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.

[illegible]



DT	24-NOV-2000	(first entry)
XX		
DE	Human oxytocin secretion promoting peptide SEQ ID NO: 36.	
KX	Human: oxytocin secretion promoter; G protein-coupled receptor protein;	
KM	treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;	
KW	caesarean section; artificial fertilization; galactostasis; goat; pig;	
RK	veterinary medicine; milk production.	
XX		
OS	Homo sapiens.	
XX		
PN	MO200038704-A1.	
PD	06-JUL-2000.	
PE	22-DEC-1999; 99WO-JP07199.	
PR	25-DEC-1998; 98JP-0369585.	
PA	(TAKE ) TAKEDA CHEM IND LTD.	
PI	Matsumoto H, Kikada C, Hinuma S;	
DR	WPI: 2000-452298/39.	
PT	Physiologically-active polypeptide recognized as ligand by G	
PT	protein-coupled receptor protein, for promoting secretion of oxytocin,	
PT	as drugs for diseases relating to oxytocin secretion and in veterinary	
PT	medicine -	
XX		
PS	Disclosure: Page 63; 72pp; Japanese.	
XX		
CC	This invention describes a novel oxytocin secretion-regulating agent	
CC	which contains a ligand peptide or its salt for the G protein-coupled	
CC	receptor protein. It is useful in the form of drugs for ameliorating,	
CC	preventing and treating diseases relating to oxytocin secretion e.g.,	
CC	weak pains and atonic bleeding, before and after expulsion of placenta,	
CC	uterine recovery failure, caesarean section, stoppage of artificial	
CC	fertilization or galactostasis and is also applicable in veterinary	
CC	medicine for promoting milk production in cow, goat and pig. This	
CC	sequence represents a human peptide which acts as an oxytocin secretion	
CC	promoter.	
SQ	Sequence 21 AA:	
XX		
Query Match	97.1%; Score 100; DB 21; Length 21;	
Best Local Similarity	89.5%; Pred. No. 1.9e-10;	
Matches 17; Conservative	0; Mismatches 2; Indels 0; Gaps 0;	
OY	1 PDINPAMYXXRGIRPVGRF 19	
Db	2 pdlnpawysrglrpvgrf 20	
RESULT 9		
AAG62535		
ID AAG62535 standard; peptide: 21 AA.		
AC AAG62535;		
XX		
DT 24-AUG-2001 (first entry)		
XX		
DE Human CRH releasing protein related peptide SEQ ID NO: 36.		
KX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;		
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;		
KW Addison's disease; adrenal gland hyperfunction; obesity.		
OS Homo sapiens.		
XX		
PN WO200135984-A1.		
XX		

```

PD 25-MAY-2001.
XX
XX PF 17-NOV-2000; 2000WO-JP08119.
XX
XX PR 18-NOV-1999; 99JP-0337900.
XX PR 26-SEP-2000; 2000JP-0297073.
XX
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX
XX PI Kitada C, Matsumoto H, Hinuma S;
XX DR WPI; 2001-355552/37.
XX
XX PT Use of G protein receptor ligand or peptide for controlling
XX corticotropin releasing hormone secretion -
XX
XX PS Disclosure; Page 75; 90pp; Japanese.
XX
XX The present sequence describes a method of controlling the secretion of
XX corticotropin releasing hormone (CRH), involving the use of a G protein
XX receptor ligand. This can be used to control the secretion of CRH and is
XX useful as an anesthetic or for treating, preventing or ameliorating
XX diseases associated with CRH secretion such as hyperaldosteronism,
XX hypercortisolism, secondary or chronic hypoadrenocorticism, Addison's
XX disease (including bordom, nausea, pigmentation, hypogonadism, hair
XX loss, and hypotension), adrenal gland hypofunction and obesity. The
XX present sequence is a peptide used in the exemplification of the
XX invention.
XX
XX SQ Sequence 21 AA;
XX
XX Query Match 97.1%; Score 100; DB 22; Length 21;
XX Best Local Similarity 89.5%; Pred. No. 1.9e-10;
XX Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX
XX QY 1 PDINPAWYXXRGIRPVGRF 19
XX ||||||| |||||||
XX 2 pdlnpawyasrglrpvgrf 20
XX
XX DB
XX
XX RESULT 10
XX AAW31396
XX ID AAW31396 standard; Peptide: 22 AA.
XX
XX AC AAW31396;
XX
XX DT 06-APR-1998 (first entry)
XX
XX DE Human type G protein-coupled receptor ligand fragment 6.
XX
XX KW G protein-coupled receptor; ligand binding; pharmaceutical;
XX modular; pituitary; central nervous system; pancreas; prophylactic;
XX therapeutic agent.
XX
XX OS Homo sapiens.
XX
XX PN WO9724436-A2.
XX
XX PD 10-JUL-1997.
XX
XX PF 26-DEC-1996; 96WO-JP03821.
XX
XX PR 18-SEP-1996; 96JP-0246573.
XX PR 28-DEC-1995; 95JP-0343371.
XX PR 15-MAR-1996; 96JP-0059419.
XX PR 12-AUG-1996; 96JP-0211805.
XX
XX PA (TAKE ) TAKEDA CHEM IND LTD.
XX
XX PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
XX PI Kawamata Y, Kitada C;
XX

```

DR WPI; 1997-363672/33.  
DR N-PSDB; AAV02433.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
XX This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosacchara. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 22 AA:  
  
Query Match 97.1%; Score 100; DB 18; Length 22;  
Best Local Similarity 89.5%; Pred. No. 2e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 PDINPAMYXXRGIRPVGRF 19  
| | | | | | | | | | | | | | | | | | | | | |  
2 pdinpwysasrgirpvgrf 20  
Db  
  
RESULT 11  
AAB10367  
ID AAB10367 standard; peptide; 22 AA.  
XX  
AC AAB10367;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PT 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-452298/39.  
PT physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,

PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
PS Disclosure; Page 64; 72pp; Japanese.  
XX  
XX This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
XX promoter.  
XX  
SQ Sequence 22 AA:  
  
Query Match 97.1%; Score 100; DB 21; Length 22;  
Best Local Similarity 89.5%; Pred. No. 2e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 PDINPAMYXXRGIRPVGRF 19  
| | | | | | | | | | | | | | | | | | | | | |  
2 pdinpwysasrgirpvgrf 20  
Db  
  
RESULT 12  
AAG62536  
ID AAG62536 standard; peptide; 22 AA.  
XX  
AC AAG62536;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 37.  
XX  
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PE 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
XX  
DR WPI; 2001-355552/37.  
XX  
PT use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Disclosure; Page 75; 90pp; Japanese.  
XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the

CC Invention.  
XX  
SQ Sequence 22 AA;

Query Match 97.1%; Score 100; DB 22; Length 22;  
Best Local Similarity 89.5%; Pred. No. 2e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 PDINPAMYXXRCIRPVGRF 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 2 pdinpamyasrgirpvgrf 20

RESULT 13  
AAW31391  
ID AAW31391 standard; Peptide: 31 AA.

XX  
AC AAW31391;  
XX  
DT 06-APR-1998 (first entry)  
XX

XX Human type G protein-coupled receptor ligand fragment 1.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
therapeutic agent.

XX Homo sapiens.

XX MO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96MO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

XX 28-DEC-1995; 95JP-0343371.

XX 15-MAR-1996; 96JP-0059419.

XX 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
PI Kawamata Y, Kitada C;

XX WPI: 1997-363672/33.

XX N-PSDB; AAV02428.

XX Claim 2; Page 184; 258pp; English.

This sequence represents a peptide fragment from a novel human type  
ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
sequence represented in AAW31390 and is used in an assay to monitor  
ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
compositions containing this ligand may be used as a pituitary function  
modulator, a central nervous system modulator or a pancreatic function  
modulator. This ligand could have specific applications as a  
prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
crania, growth hormone secretory disease, hyper- and polyphagia,  
hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
acute myocardial infarction, infertility, spinocerebellar degeneration,  
bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
oligoplasia. Assays can also be developed to screen compounds which are  
capable of altering the binding activity of the ligand affecting  
activation of the G protein-coupled receptor protein.

XX  
SQ Sequence 31 AA;

Query Match 97.1%; Score 100; DB 18; Length 31;  
Best Local Similarity 89.5%; Pred. No. 2.9e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 PDINPAMYXXRCIRPVGRF 19  
| | | | | | | | | | | | | | | | | | | | | |  
Db 13 pdinpamyasrgirpvgrf 31

RESULT 14  
AAW97235  
ID AAW97235 standard; peptide: 31 AA.

XX  
AC AAW97235;  
XX  
DT 06-MAY-1999 (first entry)  
XX

XX Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst caccogenesis;  
menopausal syndrome; euthyroid; hypometabolism; lactation;

KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
prolactinoma; infertility; impotence; amenorrhea; galactorrhea;

KW acromegaly; Chiari-Frömmel syndrome; Argon2-del castillo syndrome;  
Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KW contraceptive; placental function; choriorcarcinoma; hydatid mole;  
irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;

KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WO9858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98MO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI: 1999-105614/09.

XX Claim 3; Page 159; 241pp; English.

The present sequence represents a human type ligand fragment. It  
is used in the course of the invention. The specification describes  
an agent for modulating prolactin secretion which comprises a  
ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
protein. The agents for promoting prolactin secretion can be used for  
treating or preventing hypovarianism, gonocyst caccogenesis, menopausal  
syndrome, euthyroid or hypometabolism. They can be used for promoting  
lactation in a domestic mammal and as an aphrodisiac. The agents for  
inhibiting prolactin secretion can be used for treating or preventing  
pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
acromegaly, Chiari-Frömmel syndrome, Argon2-del castillo syndrome,  
Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
The inhibitory agents can also be used as contraceptives. The agents for  
modulating placental function can be used for treating or preventing  
choriocarcinoma, hydatid mole, irruption mole, abortion, unthrifty fetus,  
abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.





GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:57 ; Search time 136.62 Seconds  
(without alignments)  
3.576 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_2\_21  
Perfect score: 103  
Sequence: 1 PDINPAMXXRGIRPVGRFX 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCrus.COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfilest.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	100	97.1	20	3 US-09-105-678A-46	Sequence 46, Appl
2	100	97.1	20	4 US-08-776-971-64	Sequence 64, Appl
3	100	97.1	20	4 US-09-421-208-46	Sequence 46, Appl
4	100	97.1	21	3 US-09-105-678A-47	Sequence 47, Appl
5	100	97.1	21	4 US-08-776-971-65	Sequence 65, Appl
6	100	97.1	21	4 US-09-421-208-47	Sequence 47, Appl
7	100	97.1	22	3 US-09-105-678A-48	Sequence 48, Appl
8	100	97.1	22	4 US-08-776-971-66	Sequence 66, Appl
9	100	97.1	22	4 US-09-421-208-48	Sequence 48, Appl
10	100	97.1	31	3 US-09-105-678A-9	Sequence 9, Appl
11	100	97.1	31	3 US-09-105-678A-43	Sequence 43, Appl
12	100	97.1	31	4 US-08-776-971-61	Sequence 61, Appl
13	100	97.1	31	4 US-09-421-208-9	Sequence 9, Appl
14	100	97.1	31	4 US-09-421-208-43	Sequence 43, Appl
15	100	97.1	32	3 US-09-105-678A-44	Sequence 44, Appl
16	100	97.1	32	4 US-08-776-971-62	Sequence 62, Appl
17	100	97.1	32	4 US-09-421-208-44	Sequence 44, Appl
18	100	97.1	33	3 US-09-105-678A-45	Sequence 45, Appl
19	100	97.1	33	4 US-08-776-971-63	Sequence 63, Appl
20	100	97.1	33	4 US-09-421-208-45	Sequence 45, Appl
21	100	97.1	37	4 US-08-776-971-59	Sequence 59, Appl
22	100	97.1	37	4 US-08-776-971-135	Sequence 135, App
23	100	97.1	37	4 US-08-776-971-138	Sequence 138, App
24	99	96.1	20	3 US-09-105-678A-34	Sequence 34, Appl
25	99	96.1	20	3 US-09-105-678A-40	Sequence 40, Appl
26	99	96.1	20	4 US-08-776-971-8	Sequence 8, Appl
27	99	96.1	20	4 US-08-776-971-50	Sequence 50, Appl

28	99	96.1	20	4 US-08-776-971-98	Sequence 98, Appl
29	99	96.1	20	4 US-09-421-208-34	Sequence 34, Appl
30	99	96.1	20	4 US-09-421-208-40	Sequence 40, Appl
31	99	96.1	21	3 US-09-105-678A-35	Sequence 35, Appl
32	99	96.1	21	3 US-09-105-678A-41	Sequence 41, Appl
33	99	96.1	21	4 US-08-776-971-9	Sequence 9, Appl
34	99	96.1	21	4 US-08-776-971-51	Sequence 51, Appl
35	99	96.1	21	4 US-09-421-208-35	Sequence 35, Appl
36	99	96.1	21	4 US-09-421-208-41	Sequence 41, Appl
37	99	96.1	22	3 US-09-105-678A-36	Sequence 36, Appl
38	99	96.1	22	3 US-09-105-678A-42	Sequence 42, Appl
39	99	96.1	22	4 US-08-776-971-10	Sequence 10, Appl
40	99	96.1	22	4 US-08-776-971-52	Sequence 52, Appl
41	99	96.1	22	4 US-09-421-208-36	Sequence 36, Appl
42	99	96.1	22	4 US-09-421-208-42	Sequence 42, Appl
43	99	96.1	31	3 US-09-105-678A-7	Sequence 7, Appl
44	99	96.1	31	3 US-09-105-678A-8	Sequence 8, Appl
45	99	96.1	31	3 US-09-105-678A-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46  
Query Match 97.1% Score 100; DB 3; Length 20;  
Best Local Similarity 89.5% Pred. No. 1.3e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0;

QY 1 PDINPAMYXXRGIRPVGRF 19  
Db 2 PDINPAMYASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64  
; Sequence 64, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Fujii  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTED for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211605  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 97.1%; Score 100; DB 4; Length 20;  
Best Local Similarity 89.5%; Pred. No. 1.3e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDINPAMYXXRGIRPVGRF 19  
Db 2 PDINPAMYASRGIRPVGRF 20

## RESULT 3

US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 97.1%; Score 100; DB 4; Length 20;  
Best Local Similarity 89.5%; Pred. No. 1.3e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDINPAMYXXRGIRPVGRF 19  
Db 2 PDINPAMYASRGIRPVGRF 20

## RESULT 4

US-09-105-678A-47  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 97.1%; Score 100; DB 3; Length 21;  
Best Local Similarity 89.5%; Pred. No. 1.4e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
Db 2 PDINPAMYASRGIRPVGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/0876971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 97.1%; Score 100; DB 4; Length 21;  
Best Local Similarity 89.5%; Pred. No. 1.4e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
Db 2 PDINPAMYASRGIRPVGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 97.1%; Score 100; DB 4; Length 21;  
Best Local Similarity 89.5%; Pred. No. 1.4e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
||||||| |||||||  
Db 2 PDINPAMYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 97.1%; Score 100; DB 3; Length 22;  
Best Local Similarity 89.5%; Pred. No. 1.4e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
||||||| |||||||  
Db 2 PDINPAMYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 97.1%; Score 100; DB 4; Length 22;  
Best Local Similarity 89.5%; Pred. No. 1.4e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
||||||| |||||||  
Db 2 PDINPAMYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

```
Sequence 48, Application US/09421208
Patent No. 6258561
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/421,208
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 09/105,678
FILING DATE: 26-JUN-1998
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 48:
SEQUENCE CHARACTERISTICS:
LENGTH: 22 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-421-208-48

Query Match          97.1%; Score 100; DB 4; Length 22;
Best Local Similarity 89.5%; Pred. No. 1.4e-10;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 PDINPAWYXXRCIRPVGR 19
    ||||| |||||
DB 2 PDINPAWYASRCIRPVGR 20
```

```
RESULT 10
US-09-105-678A-9
Sequence 9, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
```

```
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 9:
SEQUENCE CHARACTERISTICS:
LENGTH: 31 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-09-105-678A-9
```

```
Query Match          97.1%; Score 100; DB 3; Length 31;
Best Local Similarity 89.5%; Pred. No. 2.1e-10;
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 PDINPAWYXXRCIRPVGR 19
    ||||| |||||
DB 13 PDINPAWYASRCIRPVGR 31
```

```
RESULT 11
US-09-105-678A-43
Sequence 43, Application US/09105678A
Patent No. 6103882
GENERAL INFORMATION:
APPLICANT: Suenaga, Masato
APPLICANT: Moriya, Takeo
APPLICANT: Tanaka, Yoko
APPLICANT: Nishimura, Osamu
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
NUMBER OF SEQUENCES: 52
CORRESPONDENCE ADDRESS:
ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/105,678A
FILING DATE: 26-JUN-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 172118/1997
FILING DATE: 27-JUN-1997
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 48466-342
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
```

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678a-43

Query Match 97.1%; Score 100; DB 3; Length 31;  
Best Local Similarity 89.5%; Pred. No. 2,1e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
Db 13 PDINPAMYASRGIRPVGRF 31

RESULT 12

US-08-776-971-61  
Sequence 61, Application US/08776971B  
Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji

Hadeta, Yugo

Kawamata, Yuji

Hosoya, Masaki

Fujii, Ryo

Fukushima, Shoji

Kitada, Chieko

TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE

NUMBER OF SEQUENCES: 140

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

OPERATING SYSTEM: IBM compatible

SOFTWARE: FASTSEQ for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B

FILING DATE: 06-Feb-1997

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821

FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371

FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419

FILING DATE: 15-MAR-1996

APPLICATION NUMBER: JP 8/211805

FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573

FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 47176

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 97.1%; Score 100; DB 4; Length 31;  
Best Local Similarity 89.5%; Pred. No. 2,1e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
Db 13 PDINPAMYASRGIRPVGRF 31

RESULT 13

US-09-421-208-9

Sequence 9, Application US/09421208

Patent No. 6238561

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato

APPLICANT: Moriya, Takeo

APPLICANT: Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP

STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

OPERATING SYSTEM: IBM PC compatible

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,208

FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678

FILING DATE: 26-JUN-1998

APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026

REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids

TYPE: amino acid

STRANDEDNESS:

TOPOLOGY: linear

MOLECULE TYPE: peptide

US-09-421-208-9

Query Match 97.1%; Score 100; DB 4; Length 31;  
Best Local Similarity 89.5%; Pred. No. 2,1e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGRF 19  
Db 13 PDINPAMYASRGIRPVGRF 31

RESULT 14

US-09-421-208-43  
; Sequence 43, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-43

Query Match 97.1%; Score 100; DB 4; Length 31;  
Best Local Similarity 89.5%; Pred. No. 2.1e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDINPAMYXXRGIRPVGRF 19  
||||| |||||||  
DB 13 PDINPAMYASRGIRPVGRF 31

RESULT 15  
US-09-105-678A-44  
; Sequence 44, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-44

Query Match 97.1%; Score 100; DB 3; Length 32;  
Best Local Similarity 89.5%; Pred. No. 2.2e-10;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PDINPAMYXXRGIRPVGRF 19  
||||| |||||||  
DB 13 PDINPAMYASRGIRPVGRF 31

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:23:58 ; Search time 172.41 seconds

(without alignments)  
11.147 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_2\_21

Perfect score: 103

Sequence: 1 PDINPAMYXXRGIRPVGRFX 20

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	99	96.1	83	JC7607	prolactin-releasing
2	49	47.6	790	T47959	hydroxybenzoate oc
3	46	44.7	333	H82852	hypothetical prote
4	45	43.7	767	T21959	conserved hypotnet
5	44.5	43.2	664	F83376	hypothetical prote
6	44	42.7	250	G83400	hypothetical prote
7	44	42.7	284	F71015	hypothetical prote
8	43	41.7	220	G83292	probable glutathio
9	43	41.7	232	G75608	hypothetical prote
10	43	41.7	309	T32376	hypothetical prote
11	43	41.7	376	A48197	opsin, ocellar - A
12	43	41.7	376	B48197	opsin, lateral eye
13	43	41.7	476	G64720	probable amino aci
14	43	41.7	476	G90629	probable inner mem
15	43	41.7	476	G85480	inner membrane tra
16	43	41.7	719	S61046	ARL1 protein - yea
17	42.5	41.3	443	T21499	hypothetical prote
18	42.5	41.3	1501	T45623	hypothetical prote
19	42	40.8	428	F81694	pyruvate dehydroge
20	42	40.8	433	H87660	peptidoglycan-bind
21	41.5	40.3	345	H84012	N-acetylglutamate
22	41	39.8	226	A87664	hypothetical prote
23	41	39.8	284	A75117	hypothetical prote
24	41	39.8	338	T20100	hypothetical prote
25	41	39.8	342	B64395	malic acid transpo
26	41	39.8	343	T46534	probable FMN-depen
27	41	39.8	347	H64371	malic acid transpo
28	41	39.8	476	AG0502	probable amino aci
29	41	39.8	545	A87448	conserved hypotnet

30	41	39.8	986	1	ORYGA	spect receptor p
31	41	39.8	2155	2	T30197	alpha tectorin - m
32	40	38.8	128	2	S76955	hypothetical prote
33	40	38.8	184	2	T35841	probable membrane
34	40	38.8	324	2	T35901	probable arac fami
35	40	38.8	329	2	H70744	hypothetical prote
36	40	38.8	341	2	T35426	probable oxidoredu
37	40	38.8	359	2	T40084	PMP domain protei
38	40	38.8	390	2	G82844	cysteine synthase
39	40	38.8	419	2	AH3166	hypothetical prote
40	40	38.8	430	1	B69009	conserved hypotnet
41	40	38.8	462	2	T00708	violaxanthin de-ep
42	40	38.8	468	2	C83160	nitrite extrusion
43	40	38.8	486	2	AF1174	lysine-specific pe
44	40	38.8	486	2	AG1531	lysine-specific pe
45	40	38.8	546	2	A32260	cholesterol oxidas

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
Riyamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
A:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRRP  
A:Introns: 33/1

Query Match 96.1%; Score 99; DB 2; Length 83;  
Best Local Similarity 89.5%; Pred. No. 2.3e-09;  
Matches 17; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 PDINPAMYXXRGIRPVGRF 19  
Db 34 PDINPAMYTGIRPVGRF 52  
RESULT 2  
T47959  
hypothetical protein F15G16.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Meyer, K.F.X.; O submitted to the Protein Sequence Database, January 2000  
A:Reference number: 224480  
A:Accession: T47959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEH>  
A:Cross-references: EMBL:AL132959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 39/1; 678/2; 698/3; 773/2  
A:Note: F15G16.60



RESULT 7  
 F71015  
 hypothetical protein PH1420 - Pyrococcus horikoshii  
 C:Species: Pyrococcus horikoshii  
 C:Date: 14-Aug-1998 #sequence\_revision 14-Aug-1998 #text\_change 20-Jun-2000  
 C:Accession: F71015  
 R:Kawarabayashi, Y.; Sawada, M.; Horikawa, H.; Halkawa, Y.; Hiro, Y.; Yamamoto, S.; Sekin  
 M.; Ohikuku, Y.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kushida, N.; Oguchi  
 DNA Res. 5, 55-76, 1998  
 A:Title: Complete sequence and gene organization of the genome of a hyper-thermophilic  
 A:Reference number: A71000; MUID:98344137  
 A:Accession: F71015  
 A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-284 <KAW>  
 A:Cross-references: GB:AF000006; NID:g3236133; PIDN:BAA30526.1; PID:g3257843  
 A:Experimental source: strain ON3  
 A:Note: this accession replaces an interim accession for a sequence replaced by GenBank  
 C:Genetics:  
 A:Gene: PH1420  
 C:Superfamily: Pyrococcus horikoshii hypothetical protein PH1420

Query Match 42.7%; Score 44; DB 2; Length 284;  
 Best Local Similarity 44.4%; Pred. No. 11;  
 Matches 8; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRPVGR 18  
 | | | | | | | | | | | | | | | | | | | | | |  
 DB 217 PYIEPTFALRGLELGR 234

RESULT 8  
 C83292  
 Probable glutathione S-transferase PA2821 [Imported] - Pseudomonas aeruginosa (strain PA  
 C:Species: Pseudomonas aeruginosa  
 C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001  
 C:Accession: C83292  
 R:Stover, C.K.; Plam, X.O.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
 adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Llm,  
 ; Lory, S.; Olson, M.V.  
 Nature 406, 959-964, 2000  
 A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
 A:Reference number: A82950; MUID:20437337  
 A:Accession: C83292  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-720 <STO>  
 A:Cross-references: GB:AE004709; GB:AE004091; NID:g9948904; PIDN:AAG06209.1; GSPDB:GN001  
 C:Genetics:  
 A:Experimental source: strain PA01  
 A:Gene: PA2821  
 C:Superfamily: placte glutathione transferase

Query Match 41.7%; Score 43; DB 2; Length 220;  
 Best Local Similarity 64.3%; Pred. No. 12;  
 Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 5 PAMYXXRGIRPVGR 18  
 | | | | | | | | | | | | | | | | | | | | | |  
 DB 39 PAMY--REISPLGR 50

RESULT 9  
 G75608  
 hypothetical protein - Deinococcus radiodurans (strain R1)  
 C:Species: Deinococcus radiodurans  
 C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 31-Mar-2000  
 C:Accession: G75608  
 R:White, O.; Elsen, J.A.; Heldelberg, J.F.; Hickey, E.K.; Peterson, J.D.; Dodson, R.J.;  
 M.; Shen, M.; Vamathevan, J.J.; Lam, P.; McDonald, L.; Ulterback, T.; Zalewski, C.; Ma  
 S.; Smith, H.O.; Venter, J.C.; Fraser, C.M.

Science 286, 1571-1577, 1999  
 A:Title: Genome sequence of the radioresistant bacterium Deinococcus radiodurans R1.  
 A:Reference number: A75250; MUID:20036896  
 A:Accession: G75608  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-732 <WHI>  
 A:Cross-references: GB:AE001862; GB:AE001825; NID:g6460468; PIDN:AAF12317.1; PID:g646  
 A:Experimental source: strain R1  
 C:Genetics:  
 A:Gene: DRA0132  
 A:Map position: 2

Query Match 41.7%; Score 43; DB 2; Length 232;  
 Best Local Similarity 46.7%; Pred. No. 13;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

OY 1 PDINPAMYXXRGIRP 15  
 | | | | | | | | | | | | | | | | | | | | | |  
 DB 130 PDHRAAMHLRGVLP 144

RESULT 10  
 T32376  
 hypothetical protein K10F12.4 - Caenorhabditis elegans  
 C:Species: Caenorhabditis elegans  
 C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
 C:Accession: T32376  
 R:Wohlmann, P.; Beck, C.  
 submitted to the EMBL Data Library, September 1997  
 A:Description: The sequence of C. elegans cosmid K10F12.  
 A:Reference number: Z21157  
 A:Accession: T32376  
 A:Status: preliminary; translated from GB/EMBL/DBU  
 A:Molecule type: DNA  
 A:Residues: 1-309 <WOH>  
 A:Cross-references: EMBL:AF025462; PIDN:AAB71002.1; GSPDB:GN00021; CESP:K10F12.4  
 A:Experimental source: strain Bristol N2; clone K10F12  
 C:Genetics:  
 A:Gene: CESP:K10F12.4  
 A:Map position: 3  
 A:Introns: 31/3; 123/2; 196/3; 239/1

Query Match 41.7%; Score 43; DB 2; Length 309;  
 Best Local Similarity 44.4%; Pred. No. 17;  
 Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

OY 1 PDINPAMYXXRGIRPVGR 18  
 | | | | | | | | | | | | | | | | | | | | | |  
 DB 131 PDRSPMYLPLKS--PIGR 146

RESULT 11  
 A48197  
 opsin, ocular - Atlantic horseshoe crab  
 C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
 C:Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
 C:Accession: A48197  
 R:Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
 Proc. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
 A:Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, Limulus poly  
 A:Reference number: A48197; MUID:93317641  
 A:Accession: A48197  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-376 <SMI>  
 A:Cross-references: EMBL:L03792; NID:g156644; PIDN:AAA28274.1; PID:g156645  
 C:Superfamily: vertebrate rhodopsin  
 C:Keywords: chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; re  
 F:318/Binding site: retinal (Lys) (covalent) #status predicted

Query Match 41.7%; Score 43; DB 2; Length 376;  
Best Local Similarity 46.2%; Pred. No. 21;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PDINPAMYXXRGI 13  
| : | | | | :  
Db 40 PPMNPLMYSLIGV 52

## RESULT 12

B48197  
C:Species: Limulus polyphemus (Atlantic horseshoe crab)  
C>Date: 17-Jul-1994 #sequence\_revision 17-Jul-1994 #text\_change 13-Aug-1999  
R:Smith, W.C.; Price, D.A.; Greenberg, R.M.; Battelle, B.A.  
Proc. Natl. Acad. Sci. U.S.A. 90, 6150-6154, 1993  
A:Title: Opsins from the lateral eyes and ocelli of the horseshoe crab, Limulus polyphemus  
A:Reference number: A48197; MUID:93317641  
A:Accession: B48197  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-376 <SMI>  
A:Cross-references: EMBL:L03791; NID:g156642; PID:AAA28273.1; PID:g156643  
C:Superfamily: vertebrate rhodopsin  
C:Keywords: chromoprotein; G protein-coupled receptor; lipoprotein; photoreceptor; retina  
F:318/Binding site: retinal (lys) (covalent) #status predicted

Query Match 41.7%; Score 43; DB 2; Length 376;  
Best Local Similarity 46.2%; Pred. No. 21;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PDINPAMYXXRGI 13  
| : | | | | :  
Db 40 PPMNPLMYSLIGV 52

## RESULT 13

G64720  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
R:Baltner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; C.  
A.; Rose, D.J.; Mau, B.; Shaoh, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G64720  
A>Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <BLAT>  
A:Cross-references: GB:AE000111; GB:U00096; NID:g1786181; PID:NA073118.1; PID:g1786186;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yaaJ  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein  
C:Keywords: amino acid transport; transmembrane protein  
F:10-26/Domain: transmembrane #status predicted <TM1>  
F:91-107/Domain: transmembrane #status predicted <TM2>  
F:142-158/Domain: transmembrane #status predicted <TM3>  
F:178-194/Domain: transmembrane #status predicted <TM4>  
F:208-224/Domain: transmembrane #status predicted <TM5>  
F:303-319/Domain: transmembrane #status predicted <TM6>  
F:349-365/Domain: transmembrane #status predicted <TM7>  
F:391-407/Domain: transmembrane #status predicted <TM8>  
F:414-430/Domain: transmembrane #status predicted <TM9>

Query Match 41.7%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 27;

Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

Oy 2 DIN-----PAWYXXRGI 13  
| : | | | | | :  
Db 120 DVNGQFRGPGAWYMARGL 137

## RESULT 14

G90629  
C:Species: Escherichia coli  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shingawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90629  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <HAY>  
A:Cross-references: GB:BA000007; PID:BA033430.1; PID:g13359463; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: Ecs0007  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 41.7%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 27;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

Oy 2 DIN-----PAWYXXRGI 13  
| : | | | | | :  
Db 120 DVNGQFRGPGAWYMARGL 137

## RESULT 15

G85480  
C:Species: Escherichia coli  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G85480  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
Miller, L.; Gottlieb, E.J.; Davis, N.W.; Llim, A.; Dimlanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G85480  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <STO>  
A:Cross-references: GB:AE005174; NID:g12512682; PID:NA054307.1; GSPDB:GN00145; UWGP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: yaaJ  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 41.7%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 27;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

Oy 2 DIN-----PAWYXXRGI 13  
| : | | | | | :  
Db 120 DVNGQFRGPGAWYMARGL 137

Search completed: September 13, 2002, 09:23:59  
Job time: 774 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:44 ; Search time 80.21 Seconds  
(without alignments)  
9.655 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_2\_21  
Perfect score: 103  
Sequence: 1 PDINPAMYXXRGIRPVGRFX 20

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 segs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt.40:\*  
Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match length	ID	Description
1	100	97.1	87 1	PRRP_HUMAN
2	99	96.1	83 1	PRRP_RAT
3	99	96.1	98 1	PRRP_BOVIN
4	45	43.7	676 1	EXL1_HUMAN
5	43	41.7	376 1	OPS2_LIMPO
6	43	41.7	376 1	OPS2_LIMPO
7	43	41.7	476 1	VAAJ_ECOLI
8	43	41.7	719 1	NRPI_YEAST
9	42	40.8	402 1	EX7L_STRCO
10	41.5	40.3	345 1	ARGC_BACHD
11	41	39.8	342 1	Y762_METJA
12	41	39.8	347 1	Y576_METJA
13	41	39.8	986 1	CYGR_ARBPV
14	40	38.8	329 1	Y493_MYCTU
15	40	38.8	546 1	CHOD_STRSQ
16	40	38.8	697 1	LCFF_HUMAN
17	39.5	38.3	860 1	VG12_BPBO3
18	39	37.9	149 1	EMRN_BP7
19	39	37.9	377 1	OPS1_HEMSA
20	39	37.9	377 1	OPS2_HEMSA
21	39	37.9	485 1	SYE_BACHD
22	39	37.9	622 1	PPCC_RAT
23	39	37.9	953 1	SVV_VIBCH
24	38.5	37.4	265 1	UBIE_RICPR
25	38.5	37.4	831 1	DPOL_THERL
26	38.5	37.4	877 1	CAD2_BOVIN
27	38.5	37.4	906 1	CAD2_HUMAN
28	38.5	37.4	906 1	CAD2_MOUSE
29	38.5	37.4	906 1	CAD2_MOUSE
30	38	36.9	158 1	RL15_AERPE
31	38	36.9	332 1	LYTB_MYCLE
32	38	36.9	383 1	CYCR_CHRYI
33	38	36.9	637 1	MYHA_DROME

34	38	36.9	951 1	SVV_ECOLI	P07118 escherichia
35	38	36.9	1723 1	AIM1_HUMAN	Q9Y4K1 homo sapien
36	38	36.9	2832 1	NDVB_RHME	P20471 rhizobium m
37	37.5	36.4	3083 1	POLG_ZYWR	O89330 z genome po
38	37	35.9	179 1	RK6_CUTVR	O46908 guillardia
39	37	35.9	289 1	LYTB_AQUAE	O67625 aquilex aeo
40	37	35.9	344 1	IN37_SPIOL	P23525 spinacia ol
41	37	35.9	385 1	YEIB_ECOLI	P25747 escherichia
42	37	35.9	443 1	PRSA_CAEEL	O16368 caenorhabdi
43	37	35.9	622 1	PPCC_MOUSE	O922v4 mus musculu
44	37	35.9	640 1	Y4CD_RHISN	P55386 rhizobium s
45	37	35.9	788 1	REC2_HAETN	P44408 haemophilus

## ALIGNMENTS

```
RESULT 1
PRRP_HUMAN STANDARD: PRT: 87 AA.
ID P81277;
AC 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
CN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
[2]
RN TISSUE SPECIFICITY.
RP PubMed=10498338;
RX Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sunino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
-1- FUNCTION: stimulates prolactin (PrL) release and regulates the
expression of prolactin through its receptor GPR10. May stimulate
lactotrophs directly to secrete PrL.
-1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AB015419; BAA29027.1; -
DR MIM: 602663; -
KW Hormone; Amidation; Signal.
FT SIGNAL 1 22 BY SIMILARITY.
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
FT MOD_RES 53 53 AMIDATION (G-S4 PROVIDE AMIDE GROUP).
SQ SEQUENCE 87 AA; 9639 MW; 229A2F3B50CF981B CRC64;
```



RA Mize C.A., Clines G.A., Massa H., Traak B.J., Lovett M.;  
RT "Identification and localization of the gene for ERTL, a third member  
RL of the multiple exostosins gene family.";  
RN Genome Res. 7:10-16(1997).  
[12]  
RP SEQUENCE FROM N.A.  
RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;  
RT "Mutations of the EXT genes in hereditary multiple exostososes in  
RL Chinese.";  
RN Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.  
[13]  
RP SEQUENCE FROM N.A.  
RA Mays W., Spicker N., Van Roy N., De Paeppe A., De Bouille K.,  
RA Willems P.J., Van Hul W., Versteeg R., Speleman F.;  
RT "Refined physical mapping and genomic structure of the ERTL gene.";  
RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.  
CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
CC reticulum (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U67191; AAC51141.1; -  
DR EMBL: AF083633; AAD02840.1; -  
DR EMBL: AF083623; AAD02840.1; JOINED.  
DR EMBL: AF083624; AAD02840.1; JOINED.  
DR EMBL: AF083625; AAD02840.1; JOINED.  
DR EMBL: AF083626; AAD02840.1; JOINED.  
DR EMBL: AF083627; AAD02840.1; JOINED.  
DR EMBL: AF083628; AAD02840.1; JOINED.  
DR EMBL: AF083629; AAD02840.1; JOINED.  
DR EMBL: AF083630; AAD02840.1; JOINED.  
DR EMBL: AF083631; AAD02840.1; JOINED.  
DR EMBL: AF083632; AAD02840.1; JOINED.  
DR EMBL: AF153980; AAF73172.1; -  
DR EMBL: AF153981; AAF73172.1; JOINED.  
DR MIM: 601738; -  
DR InterPro: IPR004263; Exostosin.  
DR Pfam: PF03016; Exostosin: 1.  
KM Anti-oncogene; Multigene family: Transmembrane; Signal-anchor.  
FT TRANSMEM 10 30  
FT CARBOHD 269 269  
SQ SEQUENCE 676 AA; 74673 MW; B5E006A8762E5633 CRC64;  
-----  
Query Match 43.7%; Score 45; DB 1; Length 676;  
Best Local Similarity 47.4%; Pred. No. 6; 8; Indels 0; Gaps 0;  
Matches 9; Conservative 2; Mismatches 8; Indels 0; Gaps 0;  
-----  
Qy 1 PDINPAMYXXRGIRPYGRF 19  
Db 401 PDPEFYIQQGSRPSGRF 419  
-----  
RESULT 5  
OPSL\_LIMPO STANDARD; PRT; 376 AA.  
AC P35360;  
DT 01-JUN-1994 (Rel. 29, Created)  
DT 01-JUN-1994 (Rel. 29, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Lateral eye opsin.  
OS Limulus polyphemus (Atlantic horseshoe crab).  
OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
OC Limulidae; Limulus.

OX NCBI\_TaxID=6850;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Lateral eye;  
RX MEDLINE=93317641; PubMed=8372495;  
RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.-A.;  
RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
RL Limulus polyphemus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
CC MEDIATE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
CC LINKED TO CIS-RETINAL.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
CC -1- TISSUE SPECIFICITY: LATERAL EYE.  
CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
CC BE PHOSPHORYLATED (BY SIMILARITY).  
CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 520 NM.  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
CC OPSIN SUBFAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: L03791; AAA28273.1; -  
DR EMBL: L03781; AAA02498.1; -  
DR PIR: B48197; B48197.  
DR GCRD: GCR\_0584; -  
DR InterPro: IPR000276; GPCR\_Rhodopsn.  
DR InterPro: IPR001760; Opsin.  
DR Pfam: PF00001; 7tm\_1: 1  
DR PRINTS: PR00237; GPCR\_RHODOPSN.  
DR PROSITE: PS00237; G\_PROTEIN\_RECPT\_FL\_1; 1.  
DR PROSITE: PS0262; G\_PROTEIN\_RECPT\_FL\_2; 1.  
DR PROSITE: PS00238; OPSIN; 1.  
KM Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
KW Phosphorylation; G-protein coupled receptor.  
FT DOMAIN 1 46  
FT TRANSMEM 47 71  
FT DOMAIN 72 83  
FT TRANSMEM 84 108  
FT DOMAIN 109 123  
FT TRANSMEM 124 143  
FT DOMAIN 144 162  
FT TRANSMEM 163 186  
FT DOMAIN 187 210  
FT TRANSMEM 211 238  
FT TRANSMEM 239 274  
FT TRANSMEM 275 298  
FT DOMAIN 299 306  
FT TRANSMEM 307 331  
FT DOMAIN 332 376  
FT DISULFID 120 197  
FT BINDING 318 318  
FT CARBOHD 17 17  
FT CARBOHD 193 193  
SQ SEQUENCE 376 AA; 42139 MW; CCE401766A06F26 CRC64;  
-----  
Query Match 41.7%; Score 43; DB 1; Length 376;  
Best Local Similarity 46.2%; Pred. No. 8; 5; Indels 0; Gaps 0;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
-----  
Qy 1 PDINPAMYXXRGIR 13  
Db 40 PPMNPLWYSILGV 52  
-----  
RESULT 6

OP\$2\_LIMPO STANDARD; PRT; 376 AA.  
 ID OP\$2\_LIMPO  
 AC P35361;  
 DT 01-JUN-1994 (Rel. 29, Created)  
 DT 01-JUN-1994 (Rel. 29, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Ocellular opsin.  
 OS Limulus polyphemus (Atlantic horseshoe crab).  
 OC Eukaryota; Metazoa; Arthropoda; Chelicerata; Merostomata; Xiphosura;  
 OC Limulidae; Limulus.  
 OX NCBI\_Taxid:6850;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Medial ocellii;  
 RX MEDLINE=93317641; PubMed=8327495;  
 RA Smith W.C., Price D.A., Greenberg R.M., Battelle B.A.;  
 RT "Opsins from the lateral eyes and ocelli of the horseshoe crab,  
 Limulus polyphemus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 90:6150-6154(1993).  
 CC -1- FUNCTION: VISUAL PIGMENTS ARE THE LIGHT-ABSORBING MOLECULES THAT  
 CC MEDIANE VISION. THEY CONSIST OF AN APOPROTEIN, OPSIN, COVALENTLY  
 CC LINKED TO CIS-RETINAL.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -1- TISSUE SPECIFICITY: OCELLAR CELLS; MEDIAN OCELLI.  
 CC -1- PTM: SOME OR ALL OF THE CARBOXYL-TERMINAL SER OR THR RESIDUES MAY  
 CC BE PHOSPHORYLATED (BY SIMILARITY).  
 CC -1- MISCELLANEOUS: THIS OPSIN HAS AN ABSORPTION MAXIMUM AT 530 NM.  
 CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF G-PROTEIN COUPLED RECEPTORS.  
 CC OPSIN SUBFAMILY.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1ab-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 DR EMBL; L03792; AAA28274.1; .  
 DR EMBL; L03782; AAA02499.1; .  
 DR PIR; A48197; A48197.  
 DR GCRDB; GCR 0585; .  
 DR InterPro: IPR000276; GPCR\_Rhodopsn.  
 DR InterPro: IPR001760; Opsin.  
 DR Pfam; PF00001; 7tm\_1; 1.  
 DR PRINTS; PR00237; GPCR\_RHODOPSN.  
 DR PROSITE; PS00237; G-PROTEIN\_RECEPTOR\_1; 1.  
 DR PROSITE; PS00262; G-PROTEIN\_RECEPTOR\_2; 1.  
 DR PROSITE; PS00238; Opsin; 1.  
 KW Photoreceptor; Retinal protein; Transmembrane; Glycoprotein; Vision;  
 KW Phosphorylation; G-protein coupled receptor.  
 FT DOMAIN 1 46  
 FT TRANSMEM 72 83  
 FT TRANSMEM 84 108  
 FT TRANSMEM 109 123  
 FT TRANSMEM 124 143  
 FT TRANSMEM 144 162  
 FT TRANSMEM 163 186  
 FT TRANSMEM 187 210  
 FT TRANSMEM 211 238  
 FT TRANSMEM 239 274  
 FT TRANSMEM 275 298  
 FT TRANSMEM 299 306  
 FT TRANSMEM 307 331  
 FT TRANSMEM 332 376  
 FT TRANSMEM 377 391  
 FT BINDING 318 318  
 FT CARBOHYD 17 17  
 FT CARBOHYD 193 193  
 SQ SEQUENCE 376 AA; 42111 MW; FA9647C40531CBF8 CRC64;

Query Match 41.7%; Score 43; DB 1; Length 376;  
 Best Local Similarity 46.2%; Pred. No. 8.1;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Y 1 PDINAWYXXRGI 13  
 D 40 PPMNPLWYSILGV 52  
 RESULT 7  
 ID YAAJ\_ECOLI STANDARD; PRT; 476 AA.  
 AC P30143;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative transporter yaaJ.  
 GN YAAJ OR B0007.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_Taxid:562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE=92334977; PubMed=1630901;  
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA Isono K., Mizobuchi K., Nakata A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 RT the 0.2-4 min region.";  
 RL Nucleic Acids Res. 20:3305-3308(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (potential).  
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY  
 CC (SAF). STRONG, TO H. INFLUENZAE H10183.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.1ab-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 -----  
 DR EMBL; D10483; NOT\_ANNOTATED\_CDS.  
 DR EMBL; AE000111; AAC73118.1; .  
 DR EcoGene; EG1155; yaaJ.  
 DR InterPro: IPR002293; AA\_rel\_permease\_1.  
 DR InterPro: IPR001463; Na\_ala\_symp.  
 DR Pfam; PF01235; Na\_ala\_symp; 1.  
 DR PRINTS; PR00175; NAALASMPORT.  
 DR PROSITE; PS00873; NA\_ALANINE\_SYMP; 1.  
 KW Hypothetical protein; Transmembrane; Inner membrane; Transport;  
 KW Symport; Complete proteome.  
 FT TRANSMEM 4 24  
 FT TRANSMEM 81 101  
 FT TRANSMEM 141 161  
 FT TRANSMEM 174 194  
 FT TRANSMEM 207 227  
 FT TRANSMEM 233 253  
 FT TRANSMEM 300 320  
 FT TRANSMEM 351 371  
 FT TRANSMEM 391 411



FT TRANSMEM 414 434 POTENTIAL.  
SQ SEQUENCE 476 AA: 51662 MW: 2f6EB2E12E126E63 CRC64:

Query Match  
Best Local Similarity 44.4%; Score 43; DB 1; Length 476;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 2 DIN-----PAMVXXRGI 13  
1:1 ||||| 11:  
Db 120 DVNGGFRGCGPAMVYMARGL 137

RESULT 8  
NRPI\_YEAST  
ID NRPI\_YEAST STANDARD; PRT; 719 AA.  
AC P32770; 012228;  
DT 01-OCT-1993 (Rel. 27, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Asparagine-rich protein (ARP protein).  
GN NRPI OR ARP1 OR ARP OR YDL167C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AH22;  
RX MEDLINE=93247548; PubMed=8483449;  
RA Wehner E.P., Rao E., Brendel M.;  
RT Molecular structure and genetic regulation of SFA, a gene  
RT responsible for resistance to formaldehyde in Saccharomyces  
RT cerevisiae, and characterization of its protein product.";  
RL Mol. Gen. Genet. 237:351-358(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-S288C;  
RA Pohl T.M.;  
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.  
CC -1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.  
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: X68020; CAA48159.1; -  
DR EMBL: 267750; CAA91579.1; -  
DR EMBL: 274215; CAA96741.1; -  
DR PIR: S31139; S31139.  
DR HSSP: P04170; 6RXN.  
DR SGD: S0002326; NRPI.  
DR InterPro: IPR001876; znf-RanBP.  
DR Pfam: PF00076; rrm\_1.  
DR Pfam: PF00641; zt-RanBP; 2.  
DR SMART: SM00360; RRM; 1.  
DR SMART: SM00547; znf\_RBZ; 2.  
DR PROSITE: PS00102; RRM; 1.  
DR PROSITE: PS00030; RRM\_RNP\_1; FALSE\_NEG.  
DR PROSITE: PS01358; ZF\_RANBP2\_1; 2.  
DR PROSITE: PS01199; ZF\_RANBP2\_2; 2.  
DR PROSITE: PS01199; ZF\_RANBP2\_2; 2.  
KW Nuclear protein; zinc-finger; RNA-binding; Repeat.  
FT DOMAIN 226 322  
FT ZN\_FING 355 384  
FT ZN\_FING 581 610  
FT DOMAIN 490 564  
FT CONFLICT 493 493 I -> N (IN REF. 1).

SQ SEQUENCE 719 AA: 79299 MW: ADA9BC09FD582669 CRC64:

Query Match  
Best Local Similarity 50.0%; Score 43; DB 1; Length 719;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 6 AMVXXRGIRPVG 17  
1:1 ||||| 11:  
Db 244 SMFTGYGVIRPVG 255

RESULT 9  
EX7L\_STRCO  
ID EX7L\_STRCO STANDARD; PRT; 402 AA.  
AC Q9PBM3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
DE (Exonuclease VII large subunit).  
GN XSEA OR SCR7 29C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Seeger K.J., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G.,  
RA Rajandream M.A.;  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3' or  
CC or 3' to 5'-direction to yield 5'-phosphomononucleotides.  
CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: AL391754; CAC05901.1; -  
DR InterPro: IPR003753; Exonuc\_VII\_L.  
DR InterPro: IPR002309; tRNA\_Synt\_2.  
DR Pfam: PF02601; Exonuc\_VII\_L; 1.  
DR Pfam: PF01336; tRNA\_ant1; 1.  
KW Hydroxylase; Nuclease; Exonuclease.  
SQ SEQUENCE 402 AA: 43882 MW: 145929A8372B4E08 CRC64:

Query Match  
Best Local Similarity 40.8%; Score 42; DB 1; Length 402;  
Matches 9; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

OY 5 PAMVXXRG-----IRPVG 17  
1:1 ||||| 11:  
Db 89 PAMVXXRGQLSRAAEIKPVG 109

RESULT 10  
ARGC\_BACHD  
ID ARGC\_BACHD STANDARD; PRT; 345 AA.  
AC Q9K8Y2;  
DT 01-MAR-2002 (Rel. 41, Created)

```

DT 01-MAR-2002 (Rel. 41, Last sequence update)
DE 01-MAR-2002 (Rel. 41, Last annotation update)
DE N-acetyl-gamma-glutamyl-phosphate reductase (EC 1.2.1.38) (AGPR) (N-
acetyl-glutamate semialdehyde dehydrogenase) (NAGSA dehydrogenase).
GN ARCC OR B12900.
OS Bacillus halodurans.
OC Bacteria: Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=8665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT *Complete genome sequence of the alkaliphilic bacterium Bacillus
halodurans and genomic sequence comparison with Bacillus subtilis.";
RT Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: N-acetyl-L-glutamate 5-semialdehyde + NADP(+)
+ phosphate -> N-acetyl-5-glutamyl phosphate + NADPH.
CC -1- PATHWAY: THIRD STEP IN ARGININE BIOSYNTHESIS.
CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
CC EMBL, AP001517; BAB0619.1;
DR InterPro: IPR000706; AGPR_act_site.
DR InterPro: IPR000534; Semialdh_ch.
DR Pfam: PF01118; Semialdehyde_dh; 1.
DR Pfam: PF02774; Semialdehyde_dhc; 1.
DR ProDom: PD003765; AGPR_act_site; 1.
DR PROSITE: PS01224; AGC; 1.
KW Arginine biosynthesis; Oxidoreductase; NADP; Complete proteome.
FT ACT_SITE 149 149 BY SIMILARITY.
SQ SEQUENCE 345 AA; 38188 MW; 3E9FA5D09FC68EA CRC64;

Query Match 40.38; Score 41.5; DB 1; Length 345;
Best Local Similarity 52.98; Pred. No. 13;
Matches 9; Conservative 1; Mismatches 6; Indels 1; Gaps 1;

OY 1 PDINPWWXXRGIRPVG 17
DB 106 PDVYEAWI-KR0AAPVG 121

RESULT 11
Y576_METJA STANDARD; PRT; 342 AA.
AC 058172;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0762.
GN M0762.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,

```

```

RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterlugg T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RT Science 273:1058-1073(1996).
RL -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII M0576 AND TO S.POMBE MALATE
PERMEASE (MAE1).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
CC -----
DR EMBL; U67521; AAB98753.1; -.
DR TIGR; M0762; -.
CC -1- SIMILARITY: BELONGS TO THE NAGSA DEHYDROGENASE FAMILY.
CC -----
CC Hypothetical protein; Transmembrane; Transport; Complete proteome.
FT TRANSMEM 8 28 POTENTIAL.
FT TRANSMEM 39 59 POTENTIAL.
FT TRANSMEM 79 99 POTENTIAL.
FT TRANSMEM 108 128 POTENTIAL.
FT TRANSMEM 142 162 POTENTIAL.
FT TRANSMEM 175 195 POTENTIAL.
FT TRANSMEM 207 227 POTENTIAL.
FT TRANSMEM 242 262 POTENTIAL.
FT TRANSMEM 276 296 POTENTIAL.
FT TRANSMEM 304 324 POTENTIAL.
SQ SEQUENCE 342 AA; 39534 MW; 08FEC3E2C4955D8 CRC64;

Query Match 39.88; Score 41; DB 1; Length 342;
Best Local Similarity 53.38; Pred. No. 16;
Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;

OY 3 INPWWXXRGIRPVG 17
DB 139 VNPGRW----IPVG 149

RESULT 12
Y576_METJA STANDARD; PRT; 347 AA.
AC 057996;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein M0576.
GN M0576.
OS Methanococcus jannaschii.
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;
OC Methanococcus.
OX NCBI_TaxID=2190;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;
RX MEDLINE=96337999; PubMed=8688087;
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reisch C.I.,
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodok A.,
RA Scott J.L., Geoghagen N.S.M., Weidman J.F., Fuhrmann J.L., Nguyen D.,
RA Uitterlugg T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,
RA Cotton M.D., Roberts K.M., Hurst M.A., Kalne B.P., Borodovsky M.,
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;
RT *Complete genome sequence of the methanogenic archaeon, Methanococcus
jannaschii.";
RT Science 273:1058-1073(1996).

```



CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: 277162; CAB00954.1; -  
DR TIGR: MT0513; -  
DR TubercuList: RV0493c; -  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 329 AA; 35427 MW; BE4B524750277B56 CRC64;

Query Match 38.8% Score 40; DB 1; Length 329;  
Best Local Similarity 42.9%; Pred. No. 22;  
Matches 6; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PDINPAMYXXRGIR 14  
Db 90 PAAGPAMEDIAGVR 103  
1 111: 1:1

RESULT 15  
CHOD\_STRSQ STANDARD: PRT: 546 AA.

ID CHOD\_STRSQ  
AC P12676;  
DT 01-OCT-1989 (Rel. 12, Created)  
DT 01-OCT-1989 (Rel. 12, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Cholesterol oxidase precursor (EC 1.1.3.6) (CHOD).  
GN CHOA.  
OS Streptomyces sp. (strain SA-COO).  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID:1931;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-89123081; PubMed-2914858;  
RA Ishizaki T., Hirayama N., Shinkawa H., Nimi O., Murooka Y.;  
RT "Nucleotide sequence of the gene for cholesterol oxidase from a  
RT Streptomyces sp.";  
RL J. Bacteriol. 171:596-601(1989).  
RN [2]  
RX X-RAY CRYSTALLOGRAPHY (1.5 ANGSTROMS), AND MUTAGENESIS.  
RP MEDLINE-99211873; PubMed-10194345;  
RA Yue Q.K., Kass I.J., Sampson N.S., Vrieland A.;  
RT "Crystal structure determination of cholesterol oxidase from  
RT Streptomyces and structural characterization of key active site  
RT mutants";  
RL Biochemistry 38:4277-4286(1999).  
CC -1- CATALYTIC ACTIVITY: Cholesterol + O(2) -> cholest-4-en-3-one +  
CC H(2)O(2).  
CC -1- COFACTOR: FAD.  
CC -1- PATHWAY: CHOLESTEROL METABOLISM.  
CC -1- SUBUNIT: MONOMER.  
CC -1- SUBCELLULAR LOCATION: Secreted.  
CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: M31939; AAA26719.1; -  
DR PIR: A32260; A32260.  
DR PDB: 1B4V; 06-JAN-99.  
DR PDB: 1B8S; 09-FEB-99.

DR PDB: 1CBO; 10-MAR-99.  
DR PDB: 1CC2; 11-MAR-99.  
DR InterPro: IPR001167; CHOD.  
DR InterPro: IPR000172; GMC\_oxred.  
DR Pfam: PF01319; CHOD; 1.  
DR PROSITE: PS00623; GMC\_OXRED\_1; 1.  
DR PROSITE: PS00624; GMC\_OXRED\_2; FALSE\_NEG.  
KW Oxidoreductase; Signal; Flavoprotein; FAD; 3D-structure.  
FT SIGNAL 1 42  
FT CHAIN 43 546 CHOLESTEROL OXIDASE.  
FT NR\_BIND 54 70 FAD (ADP PART) (POTENTIAL).  
FT ACT\_SITE 398 398 PROTON ACCEPTOR.  
FT ACT\_SITE 484 484  
FT MUTAGEN 398 398  
FT MUTAGEN 484 484  
SQ SEQUENCE 546 AA; 58993 MW; EF22A1FE3EA68D21 CRC64;  
H->N,O: REDUCED ACTIVITY.  
E->O: REDUCED ACTIVITY.

Query Match 38.8% Score 40; DB 1; Length 546;  
Best Local Similarity 36.8%; Pred. No. 36;  
Matches 7; Conservative 3; Mismatches 9; Indels 0; Gaps 0;

QY 1 PDINPAMYXXRGIRPVGR 19  
Db 98 PDKRSSWFKNRTEAPLGSF 116  
1 11: 1:1

Search completed: September 13, 2002, 09:30:44  
Job time: 1134 sec



```

DE  HYPOTHEICAL 87.4 KDA PROTEIN.
GN  F15G16.60.
OS  Arabidopsis thaliana (Mouse-ear cress).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC  eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX  NCBI_TaxID=3702;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA  Mayer K.F.X., Queller F., Salanoubat M.;
RL  Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN  [2]
RP  SEQUENCE FROM N.A.
RA  EU Arabidopsis sequencing project;
RL  Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AL132959; CAB71097.1; -
KW  Hypothetical protein.
SQ  SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 47.6%; Score 49; DB 10; Length 790;
Best Local Similarity 47.4%; Pred. No. 6.9;
Matches 9; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY  1 PDINPAMYXXRGIRPVGRF 19
    1 11 111111
Db  366 PPHNRTYGRSGRLQPHGRM 384

RESULT 3
O931Z7 PRELIMINARY; PRT; 420 AA.
AC  O931Z7;
DT  01-DEC-2001 (TREMBlrel. 19, Created)
DT  01-DEC-2001 (TREMBlrel. 19, Last sequence update)
DE  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  CHAIN LENGTH FACTOR-LIKE PROTEIN.
GN  AUR2B.
OS  Streptomyces aureofaciens.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC  Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX  NCBI_TaxID=1894;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  STRAIN-CM33239;
RC  Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;
RT  "Cloning and characterization of a new polyketide gene cluster in
RT  Streptomyces aureofaciens CCM3239.";
RL  Submitted (May-2001) to the EMBL/GenBank/DBJ databases.
DR  EMBL: AY033994; AAK61719.1; -
SQ  SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;

Query Match 45.1%; Score 46.5; DB 2; Length 420;
Best Local Similarity 52.6%; Pred. No. 9.4;
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

OY  6 AMYXX-----RGIRPVGRF 19
    11 11111111
Db  37 AMMAVLNGESGIRPVGRF 55

RESULT 4
O9PH76 PRELIMINARY; PRT; 333 AA.
AC  O9PH76;
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  HYDROXYBENZATE OCTAPRENYLTRANSFERASE.
GN  XP0068.
OS  Xylella fastidiosa.

```

```

OC  Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC  Xylella.
OX  NCBI_TaxID=2371;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  STRAIN-9A5C;
RC  MEDLINE=20365717; PubMed=10910347;
RX  Simpson A.J.C., Relnach F.C., Arruda P., Abreu F.A., Acencio M.,
RA  Alvarenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA  Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA  Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA  Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA  Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,
RA  Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA  Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,
RA  Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA  Ho P.L., Hohetsel J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA  Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA  Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA  Machado M.A., Madelra A.M.B.N., Madelra H.M.F., Marino C.L.,
RA  Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA  Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,
RA  Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA  Nhani A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA  de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA  Peixoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pesquero J.B.,
RA  Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA  de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA  da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A., Jr.,
RA  da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA  de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA  Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA  Zago M.A., Zatz M., Meidanis J., Setubal J.C.;
RT  "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL  Nature 406:151-159(2000).
DR  EMBL: AE003860; AAF82881.1; -
DR  InterPro: IPR000537; DB1A.
DR  Pfam: PF01040; UblA.1.
DR  PROSITE: PS00943; UblA: UNKNOWN_1.
KW  Complete proteome.
SQ  SEQUENCE 333 AA; 37931 MW; ECF3F4716C962B95 CRC64;

Query Match 44.7%; Score 46; DB 16; Length 333;
Best Local Similarity 53.3%; Pred. No. 8.9;
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY  3 INPAMYXXRGIRPVG 17
    11 11111
Db  54 LDPYWKILKRGDRPVG 68

RESULT 5
O9LGZ0 PRELIMINARY; PRT; 540 AA.
AC  O9LGZ0;
DT  01-OCT-2000 (TREMBlrel. 15, Created)
DT  01-OCT-2000 (TREMBlrel. 15, Last sequence update)
DT  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE  EST A0070209(R3722) CORRESPONDS TO A REGION OF THE PREDICTED
DE  GENE.
OS  Oryza sativa (Rice).
OC  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC  Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC  Ehrhartoideae; Oryzaceae; Oryza.
OX  NCBI_TaxID=4530;
RN  [1]
RP  SEQUENCE FROM N.A.
RA  STRAIN-CV. NIPONBARE;
RC  Sasaki T., Matsumoto T., Yamamoto K.;
RT  "Oryza sativa niponbare(ga3) genomic DNA, chromosome 1, PAC
RT  clone:p0702f03.";
RL  Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.

```

CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER  
 CC PROTEINS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +  
 CC PYROPHOSPHATE + PROTEIN N-UBIQUITYLATION.  
 CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.  
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-  
 CC THIOLESTER FORMATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
 DR EMBL; AP002481; BAA96583.1; -.  
 DR HSSP; P06104; IAYZ.  
 DR InterPro: IPR000608; UBQ\_conjugat.  
 DR Pfam; PF00179; UBQ\_conj\_1.  
 DR SMART; SM00212; UBQC; 1.  
 DR PROSITE; PSS0127; UBIQUITIN\_CONJUGAT\_2; 1.  
 DR Ligase; Ubiquitin conjugation.  
 KW SEQUENCE 540 AA; 60487 MW; 5DE1F4EEB75A86E CRC64;

Query Match 44.7%; Score 46; DB 10; Length 540;  
 Best Local Similarity 50.0%; Pred. No. 15;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 DINPAMYXXRGIRP 15  
 I: ||: ||: ||  
 Db 69 DLGVAMWVRVGRKP 82

RESULT 6  
 O20170 PRELIMINARY; PRT; 767 AA.  
 AC Q20170;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE F38EL1.7 PROTEIN.  
 GN F38EL1.7.  
 OS Caenorhabditis elegans.  
 OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Peloderinae; Caenorhabditis.  
 OX NCBI\_TaxId=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Matthews P.;  
 RL Submitted (Jan-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP MEDLINE=99069613; PubMed=9851916;  
 RX none.  
 RA "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology.";  
 RT Science 283:2012-2018(1998).  
 RL EMBL; Z68342; CAA92775.1; -.  
 DR InterPro: IPR000636; Catlon\_chan\_non\_lig.  
 DR InterPro: IPR001622; Channel\_pore\_K.  
 DR Pfam; PF00027; CNMP\_binding\_1.  
 DR Pfam; PF00520; ion\_trans\_1.  
 DR SMART; SM00100; CNMP; 1.  
 DR PROSITE; PSS00888; CNMP\_BINDING\_1; UNKNOWN\_1.  
 DR PROSITE; PSS0042; CNMP\_BINDING\_3; 1.  
 SQ SEQUENCE 767 AA; 89988 MW; F7ECF69DBBEACFC3 CRC64;

Query Match 43.7%; Score 45; DB 5; Length 767;  
 Best Local Similarity 75.0%; Pred. No. 32;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 PDINPAMY 8  
 I: |||  
 Db 747 PDVPAWY 754

RESULT 7

O911W4  
 ID O911W4 PRELIMINARY; PRT; 664 AA.  
 AC O911W4;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA2151.  
 GN PA2151.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxId=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;

RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004642; AAG05539.1; -.  
 DR InterPro: IPR001589; Actin\_act\_bind.  
 DR InterPro: IPR000461; Alpha\_amylase.  
 DR Pfam; PF00128; alpha-amylase; 1.  
 DR PROSITE; PSS00019; ACTININ\_1; UNKNOWN\_1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 664 AA; 76329 MW; 8F59FEEED54C308AD CRC64;

Query Match 43.2%; Score 44.5; DB 16; Length 664;  
 Best Local Similarity 60.0%; Pred. No. 34;  
 Matches 9; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

OY 1 PDINPAMYXXRGIRP 15  
 I: ||| I: | I:  
 Db 479 PDINP-WFLQSGRP 492

RESULT 8  
 O912F2 PRELIMINARY; PRT; 250 AA.  
 AC O912F2;  
 DT 01-MAR-2001 (TREMBLrel. 16, Created)  
 DT 01-MAR-2001 (TREMBLrel. 16, Last sequence update)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last annotation update)  
 DE HYPOTHETICAL PROTEIN PA1952.  
 GN PA1952.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxId=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
 RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,  
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
 RT opportunistic pathogen.";  
 RL Nature 406:959-964(2000).  
 DR EMBL; AE004622; AAG05340.1; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 250 AA; 25619 MW; B997F6BE28D792C2 CRC64;

```

RT      PER112,SC01, COX15, and COX11, five genes involved in the formation
RT      and function of the mitochondrial respiratory chain." ;
RL      Genomics 54:494-504(1998).
RN      [2]
RP      SEQUENCE FROM N.A.
RC      TISSUE=BRAIN;
RX      MEDLINE=96207227; PubMed=8619474;
RA      Andersson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;
RT      "A 'double adaptor' method for improved shotgun library
RT      construction." ;
RL      Anal. Biochem. 236:107-113(1996).
RN      [3]
RP      SEQUENCE FROM N.A.
RC      TISSUE=BRAIN;
RX      MEDLINE=97264341; PubMed=9110174;
RA      Yu W., Anderson B., Morley K.C., Muzny D.M., Ding Y., Liu W.,
RA      Ricafrente J.Y., Wentland M.A., Lennom G., Gibbs R.A.;
RT      "Large-scale concatenation cDNA sequencing." ;
RL      Genome Res. 7:353-358(1997).
RN      [4]
RP      SEQUENCE FROM N.A.
RA      de Ionlay P., Valnot I., Barrientos A., Gorbatyuk M., Tzagoloff A.,
RA      Benayoun E., Chetletan D., Kadhom N., Lombes A., Ogier de Baulny H.,
RA      Nisaud P., Munnich A., Rustin P., Kotig A.;
RT      "Mutations in bcs1, a mitochondrial respiratory chain assembly gene,
RT      are responsible for the complex III deficiency of patients with
RT      tubulopathy and liver failure." ;
RL      Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN      [5]
RP      SEQUENCE FROM N.A.
RC      TISSUE=RHABDOMYOSARCOMA;
RA      Strausberg R.;
RL      Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.
RN      [6]
RP      SEQUENCE FROM N.A.
RC      TISSUE=MUSCLE, RHABDOMYOSARCOMA;
RA      Strausberg R.;
RL      Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR      EMBL; AF026849; AAD08638.1; -
DR      EMBL; AF038195; AAB97365.1; -
DR      EMBL; AF346835; AAK29417.1; -
DR      EMBL; BC000416; AAH00416.1; -
DR      EMBL; BC007500; AAH07500.1; -
DR      InterPro; IPR003593; AAA.
DR      InterPro; IPR003959; AAA_subfam.
DR      Pfam; PF00004; AAA.1.
DR      SMART; SM00382; AAA.1.
KW      ATP-binding; Hypothetical protein.
SQ      SEQUENCE 419 AA; 47534 MW; 770F98BA62F2CBB8 CRC64;

Query Match      42.7%; Score 44; DB 4; Length 419;
Best Local Similarity 70.0%; Pred. No. 25;
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Oy      4 NPAWXXRXGI 13
        |||||
Db      211 NPKWTDRTGI 220

RESULT 11
O91022
ID      O91022      PRELIMINARY;      PRT:      220 AA.
AC      O91022;
DT      01-MAR-2001 (TREMBLrel. 16, Created)
DT      01-MAR-2001 (TREMBLrel. 16, Last sequence update)
DE      01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE      PROBABLE GLUTATHIONE S-TRANSFERASE.
GN      PA2821.
OS      Pseudomonas aeruginosa.
OC      Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;
OC      Pseudomonas.
OX      NCBI_TaxID=287;

```



RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 15692 / PA01;  
RX MEDLINE-20437337; PubMed-10984043;  
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,  
Stover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
Gardner R.L., Goltz L., Tolentino E., Mesibock-Wadman S., Yuan Y.,  
Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,  
Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
Reiser J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RA "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an  
opportunistic pathogen";  
RT NCBI\_TaxID=6239;  
RL Nature 406:959-964(2000).  
DR EMBL: AE004709; AAG06209.1; -  
DR InterPro: IPR004046; GST\_C.  
DR InterPro: IPR004045; GST\_N.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 220 AA; 24716 MW; 6596183BA6CA050 CRC64;

## Query Match

41.7%; Score 43; DB 16; Length 220;

Best Local Similarity 64.3%; Pred. No. 19;  
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 5 PANYXXRGIRPVGR 18  
Db 39 PAMY--REISPLGR 50

## RESULT 12

O9R221 PRELIMINARY; PRT; 232 AA.

AC O9R221;  
DT 01-MAY-2000 (TREMBLrel. 13, Created)  
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE HYPOTHEICAL 25.0 KDA PROTEIN.  
GN DRA0132.  
OS *Deinococcus radiodurans*.  
OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
OX NCBI\_TaxID=1299;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RI;  
RX MEDLINE-20036896; PubMed-10567266;  
RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
Dodson R.J., Haft D.H., Gwin M.L., Nelson W.C., Richardson D.L.,  
Moffat K.S., Qin H., Jiang L., Pamphile W., Crosby M., Shen M.,  
Vamathevan J.J., Lam P., McDonald L., Utterback T., Zalewski C.,  
Makarov K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
Ketchum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
Fraser C.M.;  
RT "Genome sequence of the radioresistant bacterium *Deinococcus*  
*radiodurans* RI";  
RL Science 286:1571-1577(1999).  
DR EMBL: AE001862; AAF12317.1; -  
DR TIGR: DRA0132; -  
DR InterPro: IPR003737; DUF158.  
DR Pfam: PF02585; DUF158; 1.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 232 AA; 24979 MW; A044FA2F38435DA7 CRC64;

## Query Match

41.7%; Score 43; DB 16; Length 232;

Best Local Similarity 46.7%; Pred. No. 20;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

Oy 1 PDINPAMYXXRGIRP 15  
Db 130 PDHRAAMHLRGLVP 144

## RESULT 13

O17234  
ID O17234 PRELIMINARY; PRT; 309 AA.  
AC O17234;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHEICAL 35.0 KDA PROTEIN.  
GN K10F12.4.  
OS *Caenorhabditis elegans*.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
OC Rhabditidae; Pelodierinae; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RX MEDLINE-99069613; PubMed-9851916;  
RA None.

RT "Genome sequence of the nematode *C. elegans*: a platform for  
investigating biology. The *C. elegans* Sequencing Consortium.";  
RL Science 282:2012-2018(1998).  
RN [2]

RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Wohlmann P., Beck C.;  
RT "The sequence of *C. elegans* cosmid K10F12.";  
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]

RP SEQUENCE FROM N.A.  
RC STRAIN-BRISTOL N2;  
RA Waterson R.;  
RT "Direct Submission.";  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF025462; AAB71002.1; -  
DR HSSP: P78417; 1EEM.  
DR InterPro: IPR004046; GST\_C.  
DR InterPro: IPR004045; GST\_N.  
DR Pfam: PF00043; GST\_C; 1.  
DR Pfam: PF02798; GST\_N; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 309 AA; 35021 MW; FA209193108AD2DC CRC64;

## Query Match

41.7%; Score 43; DB 5; Length 309;

Best Local Similarity 44.4%; Pred. No. 27;  
Matches 8; Conservative 3; Mismatches 5; Indels 2; Gaps 1;

Oy 1 PDINPAMYXXRGIRPVGR 18  
Db 131 PDHSPMYLPRKS--PIGR 146

## RESULT 14

O9C2P5 PRELIMINARY; PRT; 418 AA.

AC O9C2P5;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE 9130022019RIK PROTEIN.  
GN 9130022019RIK.  
OS *Mus musculus* (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C57BL/6J; TISSUE-EMBRYO;  
RX MEDLINE-21085660; PubMed-11217851;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,  
Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamane I.,  
Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
Kadoya K., Matsuda H.A., Ashburner M., Batilov S., Casavant T.,



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds  
(without alignments)  
5.280 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_3\_21  
Perfect score: 96  
Sequence: 1 DIMPWYXXRGIRPVGRFX 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A.Geneseq.032802:\*

```
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:*
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:*
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:*
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:*
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:*
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:*
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:*
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:*
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:*
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:*
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:*
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:*
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:*
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:*
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:*
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:*
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:*
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:*
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:*
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:*
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:*
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	96.9	20	18 AAW31394	Human type G prote
2	93	96.9	20	20 AAW97236	Human type ligand
3	93	96.9	20	21 AAB10365	Human oxytocin sec
4	93	96.9	20	21 AAY49294	19P2 ligand peptid
5	93	96.9	20	22 AAG62534	Human CRH releasin
6	93	96.9	20	22 AAB90992	Prolactin releasin
7	93	96.9	21	18 AAW31395	Human type G prote
8	93	96.9	21	21 AAB10366	Human oxytocin sec
9	93	96.9	21	22 AAG62535	Human CRH releasin
10	93	96.9	22	18 AAW31396	Human type G prote
11	93	96.9	22	21 AAB10367	Human oxytocin sec

12	93	96.9	22	22 AAG62536	Human CRH releasin
13	93	96.9	31	18 AAW31391	Human type G prote
14	93	96.9	31	20 AAW97235	Human type ligand
15	93	96.9	31	20 AAW87615	Human 19P2 ligand
16	93	96.9	31	21 AAB10362	Human oxytocin sec
17	93	96.9	31	21 AAY49291	19P2 ligand peptid
18	93	96.9	31	22 AAG62531	Human CRH releasin
19	93	96.9	31	22 AAB90991	Prolactin releasin
20	93	96.9	31	22 AAB90995	Human type G prote
21	93	96.9	32	18 AAW31392	Human type G prote
22	93	96.9	32	21 AAB10363	Human oxytocin sec
23	93	96.9	32	21 AAG62532	Human CRH releasin
24	93	96.9	33	18 AAW31393	Human type G prote
25	93	96.9	33	21 AAB10364	Human oxytocin sec
26	93	96.9	33	22 AAG62533	Human CRH releasin
27	93	96.9	87	18 AAW31390	Human type G prote
28	93	96.9	87	20 AAW97226	Human type ligand
29	93	96.9	87	21 AAB10361	Human oxytocin sec
30	93	96.9	87	22 AAG62530	Human CRH releasin
31	92	95.8	20	18 AAW31387	Rat type G protein
32	92	95.8	20	18 AAW31374	Bovine G protein-c
33	92	95.8	20	20 AAW97232	Bovine pituitary-d
34	92	95.8	20	20 AAW97234	Rat type ligand po
35	92	95.8	20	20 AAW95191	Bovine pituitary-d
36	92	95.8	20	20 AAW95175	Murine pituitary-d
37	92	95.8	20	21 AAB10350	Bovine oxytocin se
38	92	95.8	20	21 AAB10358	19P2 ligand peptid
39	92	95.8	20	21 AAY49301	19P2 ligand peptid
40	92	95.8	20	21 AAY49302	19P2 ligand peptid
41	92	95.8	20	22 AAG62519	Bovine CRH releas
42	92	95.8	20	22 AAG62527	Rat CRH releasin
43	92	95.8	20	22 AAB90994	Prolactin releasin
44	92	95.8	20	22 AAB90996	Prolactin releasin
45	92	95.8	20	22 AAB46554	Peptide PRP20 fra

## ALIGNMENTS

RESULT 1	
AAW31394	standard; Peptide: 20 AA.
AC	AAW31394;
XX	
DT	06-APR-1998 (first entry)
XX	
DE	Human type G protein-coupled receptor ligand fragment 4.
XX	
KW	G protein-coupled receptor; ligand binding; pharmaceutical;
KW	modulator; pituitary; central nervous system; pancreas; prophylactic;
KW	therapeutic agent.
XX	
OS	Homo sapiens.
XX	
PN	WO9724436-A2.
XX	
PD	10-JUL-1997.
XX	
PF	26-DEC-1996; 96WO-JP03821.
XX	
PR	18-SEP-1996; 96JP-0246573.
PR	28-DEC-1995; 95JP-0343371.
PR	15-MAR-1996; 96JP-0059419.
PR	12-AUG-1996; 96JP-0211805.
XX	
PA	(TAKEDA ) TAKEDA CHEM IND LTD.
XX	
PI	Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
PI	Kawamata Y, Kltada C;
XX	
DR	WPI, 1997-363672/33.
DR	N-PSDB; AAV02431.

XX Lligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAM97236 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. A central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC Turner's syndrome, diabetes, cancer, pancreatitis, renal disease,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosacchara. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
CC  
XX  
SQ Sequence 20 AA:  
  
Query Match 96.9%; Score 93; DB 18; Length 20;  
Best Local Similarity 88.9%; Pred. No. 6.3e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 DNPAMYXXRGIRPYGRF 18  
||||| |||||||  
Db 3 dlnpawyasrglrpygrf 20  
  
RESULT 2  
AAM97236  
ID AAM97236 standard; peptide: 20 AA.  
XX  
AC AAM97236;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Human type ligand polypeptide fragment.  
XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhoea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; chorioncarcinoma; hydatid mole;  
KW interruption; abortion; unfertilized fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.  
XX  
OS Homo sapiens.  
XX  
PN WO9858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPI: 1999-105614/09.  
XX  
XX Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 166; 241pp; English.  
XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonecyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenorrhoea, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhoea,  
CC acromegaly, Chiari-Frommel syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC chorioncarcinoma, hydatid mole, interruption, unfertilized fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
CC  
XX  
SQ Sequence 20 AA:  
  
Query Match 96.9%; Score 93; DB 20; Length 20;  
Best Local Similarity 88.9%; Pred. No. 6.3e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 DNPAMYXXRGIRPYGRF 18  
||||| |||||||  
Db 3 dlnpawyasrglrpygrf 20  
  
RESULT 3  
AAB10365  
ID AAB10365 standard; peptide: 20 AA.  
XX  
AC AAB10365;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kikada C, Hinuma S;  
XX WPI: 2000-452298/39.  
XX  
DR Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -

XX XX Disclosure: Page 63; 72pp; Japanese.  
PS XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G-protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
SQ Sequence 20 AA;  
  
Query Match 96.9%; Score 93; DB 21; Length 20;  
Best Local Similarity 88.9%; Pred. No. 6.3e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 DNPAMYXXRGIRPVGRF 18  
||||||| |||||||  
DB 3 dlnpawyasgrlrvpgrf 20  
  
RESULT 4  
AAV49294  
ID AAV49294 standard; peptide: 20 AA.  
XX  
AC AAV49294;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE 19p2 ligand peptide fragment.  
XX  
KW Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 20  
FT /note="C-terminal amide"  
XX  
PN WO960112-A1.  
XX  
PD 25-NOV-1999.  
XX  
PF 20-MAY-1999; 99WO-JP02650.  
XX  
PR 21-MAY-1998; 98JP-0140293.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-039381/03.  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
XX Disclosure: Page 26; 73pp; Japanese.  
PS  
XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAV49290-302 represent peptide fragments of the 19p2 ligand.

XX XX SQ Sequence 20 AA;  
XX  
Query Match 96.9%; Score 93; DB 21; Length 20;  
Best Local Similarity 88.9%; Pred. No. 6.3e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 DNPAMYXXRGIRPVGRF 18  
||||||| |||||||  
DB 3 dlnpawyasgrlrvpgrf 20  
  
RESULT 5  
AAG62534  
ID AAG62534 standard; peptide: 20 AA.  
XX  
AC AAG62534;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 35.  
XX  
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
XX  
DR WPI; 2001-355552/37.  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Claim 4; Page 75; 90pp; Japanese.  
XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion, such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
XX Sequence 20 AA;  
SQ  
  
Query Match 96.9%; Score 93; DB 22; Length 20;  
Best Local Similarity 88.9%; Pred. No. 6.3e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 DNPAMYXXRGIRPVGRF 18  
||||||| |||||||  
DB 3 dlnpawyasgrlrvpgrf 20  
  
RESULT 6

[illegible]

DT	24-NOV-2000	(first entry)
XX		
DE	Human oxytocin secretion promoting peptide SEQ ID NO: 36.	
XX		
KW	Human; oxytocin secretion promoter; G protein-coupled receptor protein;	
KM	treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;	
KW	caesarean section; artificial fertilization; galactostasis; goat; pig;	
XX	veterinary medicine; milk production.	
XX		
OS	Homo sapiens.	
XX		
PN	WO200038704-A1.	
PD	06-JUL-2000.	
PF	22-DEC-1999; 99WO-JP07199.	
PR	25-DEC-1998; 98JP-0369585.	
PA	(TAKE ) TAKEDA CHEM IND LTD.	
PI	Matsunoto H, Kltada C, Hinuma S:	
DR	WPI: 2000-452298/39.	
XX		
PT	Physiologically-active polypeptide recognized as ligand by G	
PT	protein-coupled receptor protein, for promoting secretion of oxytocin,	
PT	as drugs for diseases relating to oxytocin secretion and in veterinary	
PT	medicine -	
XX		
PS	Disclosure; Page 63; 72pp; Japanese.	
CC	This invention describes a novel oxytocin secretion-regulating agent	
CC	which contains a ligand peptide or its salt for the G protein-coupled	
CC	receptor protein. It is useful in the form of drugs for ameliorating,	
CC	preventing and treating diseases relating to oxytocin secretion e.g.,	
CC	weak pains and atonic bleeding, before and after expulsion of placenta,	
CC	uterine recovery failure, caesarean section, stoppage of artificial	
CC	fertilization or galactostasis and is also applicable in veterinary	
CC	medicine for promoting milk production in cow, goat and pig. This	
CC	sequence represents a human peptide which acts as an oxytocin secretion	
CC	promoter.	
XX		
SQ	Sequence 21 AA:	
Query Match	96.9%;	Score 93; DB 21; Length 21;
Best Local Similarity	88.9%;	Pred. No. 6.7e-10;
Matches 16; Conservative	0;	Mismatches 2; Indels 0; Gaps 0;
OY	1 DINDPMYXXRGTRPVGRF 18	
DB	3 dlnpawysgrlrvgrf 20	
RESULT 9		
AAG62535		
ID AAG62535 standard; peptide; 21 AA.		
AC AAG62535;		
DT 24-AUG-2001 (first entry)		
DE Human CRH releasing protein related peptide SEQ ID NO: 36.		
XX		
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;		
KM analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;		
KW Addison's disease; adrenal gland hyperfunction; obesity.		
OS Homo sapiens.		
XX		
PN WO200135984-A1.		
XX		

```
PD      25-MAY-2001.
XX
XX PF      17-NOV-2000; 2000WO-JP08119.
XX
XX PR      18-NOV-1999;   99JP-0327900.
PR      26-SEP-2000; 2000JP-0297073.
PA
XX      (TAKE ) TAKEDA CHEM IND LTD.
XX
XX PI      Kitada C, Matsumoto H, Hinuma S;
XX DR      WPI; 2001-355552/37.
XX
XX PR      Use of G protein receptor ligand or peptide for controlling
XX PT      corticotropin releasing hormone secretion -
XX PS      Disclosure; Page 75; 90pp; Japanese.
XX
XX CC      The present sequence describes a method of controlling the secretion of
CC CC      corticotrophin releasing hormone (CRH), involving the use of a G protein
CC CC      receptor ligand. This can be used to control the secretion of CRH and its
CC CC      useful as an analgesic or for treating, preventing or ameliorating
CC CC      diseases associated with CRH secretion such as hyperaldosteronism,
CC CC      hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's
CC CC      disease (including boreloma, nausea, pigmentation, hypogonadism, hair
CC CC      loss, and hypotension), adrenal gland dysfunction and obesity. The
CC CC      present sequence is a peptide used in the exemplification of the
CC CC      invention.
XX
XX SQ      Sequence    21 AA;
XX
XX Query Match          96.9%; Score 93; DB 22; Length 21;
XX Best Local Similarity 88.9%; Pred No. 6.7e-10;
XX Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0
XX
XX QY      1 DNPAMWXXXRCIRPVGRF 18
XX      ||||| |||||||
DB      3 dlnpwasyarqlrpygrf 20
XX
XX RESULT 10
XX AAWJ1396
XX ID      AAWJ1396 standard; Peptide; 22 AA.
XX
XX AC      AAWJ1396;
XX
XX DT      06-APR-1998 (first entry)
XX
XX DE      Human type G protein-coupled receptor ligand fragment 6.
XX
XX KW      G protein-coupled receptor; ligand binding; pharmaceutical;
XX KM      modulator; pituitary; central nervous system; pancreas; prophylactic;
XX KN      therapeutic agent.
XX OS      Homo sapiens.
XX
XX FX      WO9724436-A2.
XX
XX PD      10-JUL-1997.
XX
XX PF      26-DEC-1996;   96WO-JP03821.
XX
XX PR      18-SEP-1996;   96JP-0246573.
PR      28-DEC-1995;   95JP-0343371.
PR      15-MAR-1996;   96JP-0059419.
PR      12-AUG-1996;   96JP-0211805.
XX
XX PA      (TAKE ) TAKEDA CHEM IND LTD.
XX
XX FUJII R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;
XX PI      Kawamata Y, Kitada C;
```

DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02433.  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2: Page 186; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC truma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
XX  
SQ Sequence 22 AA:  
  
Query Match 96.9%; Score 93; DB 18; Length 22;  
Best Local Similarity 88.9%; Pred. No. 7e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 DINPAWYXXRGIRPYGRF 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 3 dinpawysrglrpygrf 20  
  
RESULT 11  
ID AAB10367 standard; peptide: 22 AA.  
XX  
XX AAB10367;  
XX  
XX 24-NOV-2000 (first entry)  
XX  
XX Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
DE  
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
XX Homo sapiens.  
OS  
XX WO200038704-A1.  
PN  
XX 06-JUL-2000.  
PD  
XX 22-DEC-1999; 99WO-JP07199.  
PF  
XX 25-DEC-1998; 98JP-0369585.  
PR  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
PI  
XX WPI: 2000-452298/39.  
DR  
XX Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,

PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
XX  
PS Disclosure; Page 64; 72pp; Japanese.  
XX  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
XX  
SQ Sequence 22 AA:  
  
Query Match 96.9%; Score 93; DB 21; Length 22;  
Best Local Similarity 88.9%; Pred. No. 7e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 DINPAWYXXRGIRPYGRF 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 3 dinpawysrglrpygrf 20  
  
RESULT 12  
ID AAG62536 standard; peptide: 22 AA.  
XX  
XX AAG62536;  
XX  
XX 24-AUG-2001 (first entry)  
XX  
XX Human CRH releasing protein related peptide SEQ ID NO: 37.  
DE  
XX  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX  
XX Homo sapiens.  
OS  
XX WO200135984-A1.  
PN  
XX 25-MAY-2001.  
PD  
XX 17-NOV-2000; 2000WO-JP08119.  
PF  
XX 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
PR  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Kitada C, Matsumoto H, Hinuma S;  
PI  
XX WPI: 2001-355552/37.  
DR  
XX  
XX Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
XX  
PS Disclosure; Page 75; 90pp; Japanese.  
XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including bordom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the



CC Invention.  
XX  
SQ Sequence 22 AA;

Query Match 96.9%; Score 93; DB 22; Length 22;  
Best Local Similarity 88.9%; Pred. No. 7e-10;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPVGRF 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 3 dinpawyastrgtrpvgrf 20

RESULT 13  
AAW31391  
ID AAW31391 standard; Peptide: 31 AA.

AC AAW31391;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 1.

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prolactin;  
KW therapeutic agent.

OS Homo sapiens.

XX WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX MPI: 1997-363672/33.

DR N-PSDB; AAV02428.

XX Ligand peptide for G protein-coupled receptor - acts by modulating

PT function in the central nervous system, pancreas and pituitary gland

XX Claim 2; Page 184; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC proapoptotic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligodactylia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.

XX  
SQ Sequence 31 AA;

Query Match 96.9%; Score 93; DB 18; Length 31;  
Best Local Similarity 88.9%; Pred. No. 1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPVGRF 18  
| | | | | | | | | | | | | | | | | | | | | |  
Db 14 dinpawyastrgtrpvgrf 31

RESULT 14  
AAW97235  
ID AAW97235 standard; Peptide: 31 AA.

AC AAW97235;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
KW acromegaly; Chari-Frommel syndrome; Argon-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; chorioncarcinoma; hydatid mole;  
KW Irruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PR (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

PI MPI: 1999-105614/09.

DR Use of G protein-coupled receptor ligands - for modulating prolactin

XX secretion or placental function, e.g. for treating menopausal

PT syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3; Page 159; 241pp; English.

CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
CC acromegaly, Chari-Frommel syndrome, Argon-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC chorioncarcinoma, hydatid mole, Irruption mole, abortion, unfertilized fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 31 AA;

Query Match 96.9%; Score 93; DB 20; Length 31;  
Best Local Similarity 88.9%; Pred. NO. 1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMXXXXGIRPVGRF 18  
||||| |||||||  
DB 14 dlnpawysrglrpvgrf 31

RESULT 15  
AAM87615  
ID AAM87615 standard; Peptide: 31 AA.

XX AAM87615;

XX 29-MAR-1999 (first entry)

XX Human 19P2 ligand.

XX 19P2 ligand; G protein coupled receptor; pituitary;

KW prolactin releasing peptide; human; dementia; breast cancer;  
therapy.

XX Homo sapiens.

XX EP887417-A2.

XX 30-DEC-1998.

XX 25-JUN-1998; 98EP-0111725.

XX 27-JUN-1997; 97JP-0172118.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;

XX WPI: 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
PT of a fusion protein, useful for preventing and treating dementia,  
PT breast cancer, renal failure and autoimmune disease

XX Claim 5; Page 35; 56pp; English.

XX This is the amino acid sequence of the human pituitary G  
CC protein-coupled receptor ligand 19P2L. A method suitable for  
CC commercial high-level production of 19P2L comprises expressing  
CC the ligand in host cells as a recombinant fusion protein e.g. with  
CC human basic fibroblast growth factor (see AAV83796-97) that has  
CC been modified to include an N-terminal cysteine residue. The  
CC ligand is released from the fusion by cyanylation followed by  
CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at  
CC high doses) prolactin secretion-inhibiting properties. It can be  
CC used in the treatment and prevention of various diseases including:  
CC senile dementia, cerebrovascular dementia, and dementia associated  
CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's  
CC disease, Pick's disease, Huntington's disease), infectious diseases  
CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
CC intoxication by drugs, metal and organic compounds), tumorigenic  
CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
CC subarachnoid haemorrhage), and other types of dementia, depression,  
CC hyperactive child syndrome (microencephalopathy) and disturbance of  
CC consciousness. It is also useful for prevention and treatment of  
CC diseases associated with prolactin hypo and hypersecretion  
CC respectively, including: hyperprolactinemia, pituitary adenoma,  
CC breast cancer, infertility, impotence and autoimmune disease  
CC (hypersecretion disorders), and seminal vesicle hypoplasia,

CC osteoporosis, menopausal syndrome and renal failure (hypersecretion  
CC disorders). The 19P2 polypeptide/amide is also useful as a test  
CC reagent for study of the prolactin secretory function or as a  
CC lactagogue in mammalian farm animals.

XX Sequence 31 AA;

Query Match 96.9%; Score 93; DB 20; Length 31;  
Best Local Similarity 88.9%; Pred. NO. 1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMXXXXGIRPVGRF 18  
||||| |||||||  
DB 14 dlnpawysrglrpvgrf 31

Search completed: September 13, 2002, 09:18:35  
Job time: 500 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 Seconds  
(without alignments)  
3.397 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_3\_21  
Perfect score: 96  
Sequence: 1 D1NPAWYXXRGIRPYGRFX 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents\_AA:\*  
1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/prodata/2/1aa/POCTUS\_COMB.pep:\*  
6: /cgn2\_6/prodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	96.9	20	3	US-09-105-678A-46
2	93	96.9	20	4	US-08-776-971-64
3	93	96.9	20	4	US-09-421-208-46
4	93	96.9	21	3	US-09-105-678A-47
5	93	96.9	21	4	US-08-776-971-65
6	93	96.9	21	4	US-09-421-208-47
7	93	96.9	22	3	US-09-105-678A-48
8	93	96.9	22	4	US-08-776-971-66
9	93	96.9	22	4	US-09-421-208-48
10	93	96.9	31	3	US-09-105-678A-9
11	93	96.9	31	4	US-09-105-678A-43
12	93	96.9	31	4	US-08-776-971-61
13	93	96.9	31	4	US-09-421-208-9
14	93	96.9	31	4	US-09-421-208-43
15	93	96.9	32	3	US-09-105-678A-44
16	93	96.9	32	4	US-08-776-971-62
17	93	96.9	32	4	US-09-421-208-44
18	93	96.9	33	3	US-09-105-678A-45
19	93	96.9	33	4	US-08-776-971-63
20	93	96.9	33	4	US-09-421-208-45
21	93	96.9	87	4	US-08-776-971-59
22	93	96.9	87	4	US-08-776-971-135
23	93	96.9	87	4	US-08-776-971-138
24	92	95.8	20	3	US-09-105-678A-34
25	92	95.8	20	3	US-09-105-678A-40
26	92	95.8	20	4	US-08-776-971-8
27	92	95.8	20	4	US-08-776-971-50

28	92	95.8	20	4	US-08-776-971-98	Sequence 98, Appl
29	92	95.8	20	4	US-09-421-208-34	Sequence 34, Appl
30	92	95.8	20	4	US-09-421-208-40	Sequence 40, Appl
31	92	95.8	21	3	US-09-105-678A-35	Sequence 35, Appl
32	92	95.8	21	3	US-09-105-678A-41	Sequence 41, Appl
33	92	95.8	21	4	US-08-776-971-9	Sequence 9, Appl1
34	92	95.8	21	4	US-08-776-971-51	Sequence 51, Appl
35	92	95.8	21	4	US-09-421-208-35	Sequence 35, Appl
36	92	95.8	21	4	US-09-421-208-41	Sequence 41, Appl
37	92	95.8	22	3	US-09-105-678A-36	Sequence 36, Appl
38	92	95.8	22	3	US-09-105-678A-42	Sequence 42, Appl
39	92	95.8	22	4	US-08-776-971-10	Sequence 10, Appl
40	92	95.8	22	4	US-08-776-971-52	Sequence 52, Appl
41	92	95.8	22	4	US-09-421-208-36	Sequence 36, Appl
42	92	95.8	22	4	US-09-421-208-42	Sequence 42, Appl
43	92	95.8	31	3	US-09-105-678A-7	Sequence 7, Appl1
44	92	95.8	31	3	US-09-105-678A-8	Sequence 8, Appl1
45	92	95.8	31	3	US-09-105-678A-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
; Sequence 46, Application us/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-46

Query Match 96.9%; Score 93; DB 3; Length 20;  
Best Local Similarity 88.9%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPVGRF 18  
||||| |||||||  
DB 3 DINPAMYASRGIRPVGRF 20

RESULT 2  
US-08-776-971-64

; Sequence 64, Application US/08776971B  
; Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Taji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS

SOFTWARE: PASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 64:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 96.9%; Score 93; DB 4; Length 20;

Best Local Similarity 88.9%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPVGRF 18  
||||| |||||||  
DB 3 DINPAMYASRGIRPVGRF 20

RESULT 3  
US-09-421-208-46

; Sequence 46, Application US/09421208  
; Patent No. 6258561

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 46:

SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 96.9%; Score 93; DB 4; Length 20;

Best Local Similarity 88.9%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPVGRF 18  
||||| |||||||  
DB 3 DINPAMYASRGIRPVGRF 20

RESULT 4

US-09-105-678A-47

; Sequence 47, Application US/09105678A  
; Patent No. 6103882

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 96.9%; Score 93; DB 3; Length 21;  
Best Local Similarity 88.9%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 D1NPAMYXXRGIRPVGRF 18  
||||| |||||||  
DB 3 D1NPAMYASRGIRPVGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Taji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 96.9%; Score 93; DB 4; Length 21;  
Best Local Similarity 88.9%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 D1NPAMYXXRGIRPVGRF 18  
||||| |||||||  
DB 3 D1NPAMYASRGIRPVGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 96.9%; Score 93; DB 4; Length 21;  
Best Local Similarity 88.9%; Pred. No. 1.3e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMVXXRGIRPVGRF 18  
||||| |||||||  
DB 3 DINPAMVYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882

GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 1992 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 96.9%; Score 93; DB 3; Length 22;  
Best Local Similarity 88.9%; Pred. No. 1.4e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMVXXRGIRPVGRF 18  
||||| |||||||  
DB 3 DINPAMVYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984

GENERAL INFORMATION:  
APPLICANT: Hnuna, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 96.9%; Score 93; DB 4; Length 22;  
Best Local Similarity 88.9%; Pred. No. 1.4e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMVXXRGIRPVGRF 18  
||||| |||||||  
DB 3 DINPAMVYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

```
; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-48

Query Match          96.9%; Score 93; DB 4; Length 22;
Best Local Similarity 88.9%; Pred. No. 1.4e-09;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DINDPAMYXXRGIRPVGRF 18
    ||||| |||||
DB 3 DINDPAMYASRGIRPVGRF 20

RESULT 10
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
;
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-9

Query Match          96.9%; Score 93; DB 3; Length 31;
Best Local Similarity 88.9%; Pred. No. 2.1e-09;
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DINDPAMYXXRGIRPVGRF 18
    ||||| |||||
DB 14 DINDPAMYASRGIRPVGRF 31

RESULT 11
US-09-105-678A-43
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIORITY APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
```

;; INFORMATION FOR SEQ ID NO: 43:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 31 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
;; MOLECULE TYPE: peptide  
US-09-105-678a-43

Query Match 96.9%; Score 93; DB 3; Length 31;  
Best Local Similarity 88.9%; Pred. No. 2.1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPYGRF 18  
Db 14 DINPAMYASRGIRPYGRF 31

RESULT 12  
US-08-776-971-61  
; Sequence 61, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776, 971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 96.9%; Score 93; DB 4; Length 31;  
Best Local Similarity 88.9%; Pred. No. 2.1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPYGRF 18  
Db 14 DINPAMYASRGIRPYGRF 31

RESULT 13  
US-09-421-208-9  
; Sequence 9, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 96.9%; Score 93; DB 4; Length 31;  
Best Local Similarity 88.9%; Pred. No. 2.1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 DINPAMYXXRGIRPYGRF 18  
Db 14 DINPAMYASRGIRPYGRF 31

RESULT 14



US-09-421-208-43  
; Sequence 43, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-43

Query Match 96.9%; Score 93; DB 4; Length 31;  
Best Local Similarity 88.9%; Pred. No. 2.1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DINPAMYXXRGIRPVGRF 18  
DB 14 DINPAMYASRGIRPVGRF 31

RESULT 15  
US-09-105-678A-44  
; Sequence 44, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-44

Query Match 96.9%; Score 93; DB 3; Length 32;  
Best Local Similarity 88.9%; Pred. No. 2.1e-09;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 DINPAMYXXRGIRPVGRF 18  
DB 14 DINPAMYASRGIRPVGRF 31

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein & protein search, using sw model

Run on: September 13, 2002, 09:23:59 : Search time 172.41 Seconds  
(without alignments)  
10.589 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_3\_21  
Perfect score: 96  
Sequence: 1 DINDPAMYXXRGIRPVGRFX 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 segs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: pIR1:\*  
2: pIR2:\*  
3: pIR3:\*  
4: pIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	92	95.8	83	2 JC7607	prolactin-releasing
2	46	47.9	333	2 H82852	hydroxybenzoate oc
3	46	47.9	790	2 T47959	hypothetical prote
4	43	44.8	220	2 C83292	probable glutathio
5	43	44.8	476	2 G64720	probable amino aci
6	43	44.8	476	2 G90629	probable inner mem
7	43	44.8	476	2 G85480	inner membrane tra
8	43	44.8	719	2 S61046	ARPI protein - yea
9	42.5	44.3	443	2 T21499	hypothetical prote
10	42	43.8	433	2 H87660	peptidoglycan-bind
11	41	42.7	226	2 AB7664	hypothetical prote
12	41	42.7	338	2 T20100	hypothetical prote
13	41	42.7	342	2 B64395	malic acid transpo
14	41	42.7	347	2 H64371	malic acid transpo
15	41	42.7	476	2 AG0502	conserved amino-aci
16	41	42.7	545	2 AB7448	conserved hypothet
17	40	41.7	184	2 T35841	probable membrane
18	40	41.7	284	2 F71015	hypothetical prote
19	40	41.7	324	2 T35901	probable arac faml
20	40	41.7	390	2 G82844	cysteine synthase
21	40	41.7	430	1 B69009	conserved hypothet
22	40	41.7	462	2 T00708	violaxanthin de-ep
23	40	41.7	486	2 AF1174	lysine-specific pe
24	40	41.7	486	2 AG1531	lysine-specific pe
25	39.5	41.1	965	2 AE0418	valine--tRNA ligas
26	39.5	41.1	779	2 T49717	related to BCS1 pr
27	39	40.6	105	2 D81319	hypothetical prote
28	39	40.6	149	1 NEB337	endodeoxyribonucle
29	39	40.6	250	2 G83400	hypothetical prote

30	39	40.6	256	2 E95936	hypothetical prote
31	39	40.6	340	2 T20102	hypothetical prote
32	39	40.6	485	2 E83663	glutamyl-tRNA synt
33	39	40.6	938	2 C84480	hypothetical prote
34	39	40.6	938	2 E82068	valyl-tRNA synthet
35	39	40.6	1296	2 T16859	hypothetical prote
36	38.5	40.1	154	2 AH0264	conserved hypothet
37	38.5	40.1	374	2 G70947	hypothetical prote
38	38.5	40.1	831	2 S26675	DNA-directed DNA p
39	38.5	40.1	877	1 IJH0CN	N-cadherin precurs
40	38.5	40.1	906	1 IJH0CN	N-cadherin 2 precurs
41	38.5	40.1	906	1 IJH0CN	N-cadherin precurs
42	38	39.6	127	2 B83157	hypothetical prote
43	38	39.6	158	2 F72725	probable ribosomal
44	38	39.6	167	2 AB2796	acetyltransferase
45	38	39.6	167	2 B97575	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Blochem. Biophys. Res. Commun. 281, 53-56, 2001  
A>Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MID:21092785; PMID:11176959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PrRP  
A:Introns: 33/1

Query Match 95.8%; Score 92; DB 2; Length 83;  
Best local Similarity 88.9%; Pred. No. 1.4e-08;  
Matches 16; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 DINDPAMYXXRGIRPVGRF 18  
Db 35 DINDPAMWTGGRIPVGRF 52  
H82852  
hydroxybenzoate octaprenyltransferase XF0068 [Imported] - Xylella fastidiosa (strain C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: H82852  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq Nature 406, 151-157, 2000  
A>Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MID:20365717  
A>Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82852  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <SIM>  
A:Cross-references: GB:AE003860; GB:AE003849; NID:99104830; PIDN:AAF82881.1; GSPDB:GN A:Experimental source: strain 945c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Accencio, M.; Alvarenga, R. Birones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer as-Neto, E.; Docena, C.; El-Dorri, H.; Facincani, A.P.; Ferrelha, A.U.S.  
submitted to Genbank, June 2000

A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigt  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B  
A:Authors: Martins, E.M.F.; Matukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmler, D.J.  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0068  
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 47.9%; Score 46; DB 2; Length 333;  
Best Local Similarity 53.3%; Pred. No. 3.8;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPAMYXXRGIRPVG 16  
DB 54 IDPYMKLARGDRPVG 68

RESULT 3  
T47959  
hypothetical protein F15G16.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; Que  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z24480  
A:Accession: T47959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEK>  
A:Cross-references: EMBL:AL132959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 39/1; 678/2; 698/3; 773/2  
A:Note: F15G16.60

Query Match 47.9%; Score 46; DB 2; Length 790;  
Best Local Similarity 50.0%; Pred. No. 9;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 3 NPAMYXXRGIRPVG 18  
DB 369 NPRTYSGRLQPHGRW 384

RESULT 4  
C83292  
probable glutathione S-transferase PA2821 [imported] - Pseudomonas aeruginosa (strain PA  
C:Species: Pseudomonas aeruginosa  
C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001  
C:Accession: C83292  
R:Stover, C.K.; Pham, X.O.; Ertwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; Br  
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; Lim,  
.; Lory, S.; Olson, M.V.  
Nature 406, 959-964, 2000  
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic patho  
A:Reference number: A82950; MUID:20437337  
A:Accession: C83292  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-220 <STO>  
A:Cross-references: GB:AE004709; GB:AE004091; NID:g9948904; PIDN:AG06209.1; GSPDB:GN001  
A:Experimental source: strain PA01  
C:Genetics:

A:Gene: PA2821  
C:Superfamily: plalce glutathione transferase

Query Match 44.8%; Score 43; DB 2; Length 220;  
Best Local Similarity 64.3%; Pred. No. 8.2;  
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 4 PAMYXXRGIRPVG 17  
DB 39 PAMY--REISPRGR 50

RESULT 5  
G64720  
probable amino acid transport protein yaaU, sodium-dependent - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
C:Accession: G64720  
R:Blatner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G64720  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <BLAT>  
A:Cross-references: GB:AE000111; GB:U00096; NID:g1786181; PIDN:ACC73118.1; PID:g17861  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yaaU  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein  
C:Keywords: amino acid transport; transmembrane protein  
F:10-26/Domain: transmembrane #status predicted <TM1>  
F:91-107/Domain: transmembrane #status predicted <TM2>  
F:142-158/Domain: transmembrane #status predicted <TM3>  
F:178-194/Domain: transmembrane #status predicted <TM4>  
F:208-224/Domain: transmembrane #status predicted <TM5>  
F:303-319/Domain: transmembrane #status predicted <TM6>  
F:349-365/Domain: transmembrane #status predicted <TM7>  
F:391-407/Domain: transmembrane #status predicted <TM8>  
F:414-430/Domain: transmembrane #status predicted <TM9>

Query Match 44.8%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 18;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 1 DIN-----PAMYXXRGI 12  
DB 120 DVNGQFRGSPAMYMARGL 137

RESULT 6  
G90629  
probable inner membrane transport protein Ecs0007 [imported] - Escherichia coli (stra  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: G90629  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and 9  
A:Reference number: A96629; MUID:21156231; PMID:11258796  
A:Accession: G90629  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA83430.1; PID:g13359463; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: Ecs0007

C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 44.8%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 18;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 1 DIN-----PAWYXXRGI 12  
Db 120 DVNGQFRGSPAWYMARGL 137

RESULT 7  
G85480  
Inner membrane transport protein [imported] - Escherichia coli (strain O157:H7, substrat  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G85480  
R:Perera, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew  
Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimmlanta, E.; Potamouzis, K.; Apodaca,  
Mature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MID:21074935; PMID:11206551  
A:Accession: G85480  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <STO>  
A:Cross-references: GB:AE005174; NID:g12512682; PIDN:AG54307.1; GSPDB:GN00145; UMG:200  
A:Experimental source: strain O157:H7, substrain EDL533  
C:Genetics:  
A:Gene: yaaJ  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 44.8%; Score 43; DB 2; Length 476;  
Best Local Similarity 44.4%; Pred. No. 18;  
Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;

OY 1 DIN-----PAWYXXRGI 12  
Db 120 DVNGQFRGSPAWYMARGL 137

RESULT 8  
S61046  
ARPI protein - Yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein D1478; protein YDL167c  
C:Species: Saccharomyces cerevisiae  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 29-Oct-1999  
C:Accession: S61046; S31139; S67719  
R:Pohl, T.M.  
Submitted to the EMBL Data Library, November 1995  
A:Reference number: S61010  
A:Accession: S61046  
A:Molecule type: DNA  
A:Residues: 1-719 <POH>  
A:Cross-references: EMBL:267750; NID:g1061256; PIDN:CAA91579.1; PID:g1061272  
R:Wehner, E.P.; Rao, E.; Brendel, M.  
Mol. Gen. Genet. 237, 351-358, 1993  
A:Title: Molecular structure and genetic regulation of SFA, a gene responsible for resis  
A:Reference number: S31136; MID:93247348  
A:Accession: S31139  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-492, 'N', 494-719 <WEH>  
A:Cross-references: EMBL:X68020; NID:g577609; PIDN:CAA48159.1; PID:g288590  
A:Note: The nucleotide sequence was submitted to the EMBL Data Library, August 1992  
R:Pohl, T.M.  
Submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67708  
A:Accession: S67719  
A:Molecule type: DNA  
A:Residues: 1-719 <POW>

A:Cross-references: EMBL:274215; NID:g1431265; PIDN:CAA98741.1; PID:e253076; PID:g143  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:NRPI; ARPI  
A:Cross-references: MIPS:YDL167c; SGD:S0002326  
A:Map position: 4L

Query Match 44.8%; Score 43; DB 2; Length 719;  
Best Local Similarity 50.0%; Pred. No. 27;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 5 AWYXXRGIRPVG 16  
Db 244 SWFTQYGVIRPVG 255

RESULT 9  
T21499  
hypothetical protein F28D1.8 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T21499  
R:Baynes, C.  
Submitted to the EMBL Data Library, April 1996  
A:Reference number: Z19430  
A:Accession: T21499  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-443 <WIL>  
A:Cross-references: EMBL:270684; PIDN:CAA94603.1; GSPDB:GN00022; CESP:F28D1.8  
A:Experimental source: clone F28D1  
C:Genetics:  
A:Gene: CESP:F28D1.8  
A:Map position: 4  
A:introns: 71/1; 103/3; 162/3; 215/2; 360/2

Query Match 44.3%; Score 42.5; DB 2; Length 443;  
Best Local Similarity 37.5%; Pred. No. 20;  
Matches 9; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

OY 2 INPAMYXXR-----GIRPVG 16  
Db 53 INPSWFDWRYSVSDGHLGIRPVG 76

RESULT 10  
H87660  
peptidoglycan-binding protein, probable [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87660  
R:Nierman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; J. Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MID:21173698; PMID:11255647  
A:Accession: H87660  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-433 <STO>  
A:Cross-references: GB:AE005673; NID:g13425020; PIDN:AAK25284.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3322

Query Match 43.8%; Score 42; DB 2; Length 433;  
Best Local Similarity 60.0%; Pred. No. 24;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 PAWYXXRGIR 13



A:Gene: SRY0006  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 42.7%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 39;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 4 PAWYXXRGI 12  
||| ||:  
Db 129 PAWYMARGL 137

Search completed: September 13, 2002, 09:23:59  
Job time: 774 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:44 ; Search time 80.21 Seconds  
(without alignments)  
9.172 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_3\_21  
Perfect score: 96  
Sequence: 1 DINDPAMYXXRGIRPVGRFX 19

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	93	96.9	87	1	PRRP_HUMAN
2	92	95.8	83	1	PRRP_RAT
3	92	95.8	98	1	PRRP_BOVIN
4	44	45.8	676	1	EXL1_HUMAN
5	43	44.8	476	1	YAAJ_ECOLI
6	43	44.8	719	1	NRP1_YEAST
7	42	43.8	402	1	EX7L_STRCO
8	41	42.7	347	1	Y762_METUA
9	41	42.7	347	1	Y76_METUA
10	39	40.6	149	1	ENRN_BPT7
11	39	40.6	485	1	SYE_BACHD
12	39	40.6	953	1	SVY_VIRCH
13	38.5	40.1	831	1	DPO1_THEFL
14	38.5	40.1	877	1	CAD2_BOVIN
15	38.5	40.1	906	1	CAD2_HUMAN
16	38.5	40.1	906	1	CAD2_MOUSE
17	38.5	40.1	906	1	CAD2_RAT
18	38.5	40.1	906	1	RL15_AERPE
19	38	39.6	332	1	LYTB_MYCLE
20	38	39.6	637	1	MTHA_DROME
21	38	39.6	951	1	SYV_ECOLI
22	38	39.6	951	1	SYV_ECOLI
23	38	39.6	1723	1	AIM1_HUMAN
24	37	38.5	179	1	RY6_GUTTH
25	37	38.5	289	1	LYTB_AQUAE
26	37	38.5	344	1	IN37_SPIOL
27	37	38.5	376	1	OP5L_LIMPO
28	37	38.5	376	1	OP52_LIMPO
29	37	38.5	385	1	YE1B_ECOLI
30	37	38.5	640	1	Y4CD_RHISN
31	37	38.5	788	1	REC2_HAEIN
32	37	38.5	987	1	EPH4_MOUSE
33	36	37.5	81	1	RL31_SYNY3

34	36	37.5	204	1	TRPF_SULSO
35	36	37.5	241	1	Y63D_YEAST
36	36	37.5	329	1	Y493_MYCTU
37	36	37.5	334	1	YNHG_ECOLI
38	36	37.5	459	1	VL2_HPV6A
39	36	37.5	459	1	VL2_HPV6B
40	36	37.5	557	1	ALC_CANAL
41	36	37.5	596	1	SDP_EIMBO
42	36	37.5	649	1	RA32_SCHPO
43	36	37.5	987	1	EPH4_HUMAN
44	36	37.5	994	1	SYV_XYUFA
45	36	37.5	1282	1	TP2M_DICDI

## ALIGNMENTS

RESULT	ID	PRRP_HUMAN	STANDARD	PRT	87 AA.
AC	PRRP_HUMAN	PRRP_HUMAN	STANDARD	PRT	87 AA.
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP3; Proactin-releasing peptide PrRP20].				
DE	hormone) [Contains: Proactin-releasing peptide PrRP3; Proactin-releasing peptide PrRP20].				
GN	PRP.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=98268781; PubMed=9607765;				
RA	Hiduma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;				
RA	"A proactin-releasing peptide in the brain.";				
RT	Nature 393:272-276(1998).				
RT	(2)				
RP	TISSUE SPECIFICITY.				
RX	PubMed=10498338;				
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hiduma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;				
RA	"Tissue distribution of proactin-releasing peptide (PrRP) and its receptor.";				
RT	Regul. Pept. 83:1-10(1999).				
RL	-1- FUNCTION: Stimulates proactin (PRL) release and regulates the expression of proactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.				
CC	-1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: AB015419; BAA29027.1; -				
DR	MIM: 602663; -				
KW	Hormone; Amidation; Signal.				
FT	SIGNAL	1	22		
FT	PEPTIDE	23	53		
FT	PEPTIDE	34	53		
FT	MOD. RES	53	53		
SO	SEQUENCE	87 AA;	9639 MW;		

BY SIMILARITY.  
PROACTIN-RELEASING PEPTIDE PRRP3.  
PROACTIN-RELEASING PEPTIDE PRRP20.  
AMIDATION (G-54 PROVIDE AMIDE GROUP).



RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
 RT "Identification and localization of the gene for EXT1, a third member  
 of the multiple exostoses gene family.";  
 RL Genome Res. 7:10-16(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;  
 RT "Mutations of the EXT genes in hereditary multiple exostoses in  
 Chinese.";  
 RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Wuyts W., Spieker N., Van Roy N., De Paeppe A., De Bouille K.,  
 RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;  
 RT "Refined physical mapping and genomic structure of the EXT1 gene.";  
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U67191; AAC51141.1; -  
 DR EMBL: AF083633; AAD02840.1; -  
 DR EMBL: AF083623; AAD02840.1; JOINED.  
 DR EMBL: AF083624; AAD02840.1; JOINED.  
 DR EMBL: AF083625; AAD02840.1; JOINED.  
 DR EMBL: AF083626; AAD02840.1; JOINED.  
 DR EMBL: AF083627; AAD02840.1; JOINED.  
 DR EMBL: AF083628; AAD02840.1; JOINED.  
 DR EMBL: AF083629; AAD02840.1; JOINED.  
 DR EMBL: AF083630; AAD02840.1; JOINED.  
 DR EMBL: AF083631; AAD02840.1; JOINED.  
 DR EMBL: AF083632; AAD02840.1; JOINED.  
 DR EMBL: AF153980; AAF73172.1; -  
 DR EMBL: AF153991; AAF73172.1; JOINED.  
 DR MIM: 601738; -  
 DR InterPro: IPR004263; Exostosin.  
 DR Pfam: PF03016; Exostosin: 1.  
 KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.  
 FT TRANSMEM 10 30  
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT CARBOHYD 269 269  
 FT N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT SEQUENCE 676 AA; 74673 MW; B5E06A8762E5633 CRC64;  
 Query Match 45.8%; Score 44; DB 1; Length 676;  
 Best Local Similarity 53.3%; Pred. No. 5;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 OY 4 PAMVXXRGIRPYGRF 18  
 DB 405 PYYIQQGSGRPGRF 419  
 RESULT 5  
 ID YAAI\_ECOLI STANDARD; PRT; 476 AA.  
 AC P30143;  
 DT 01-JUL-1993 (Rel. 26, Created)  
 DT 01-JUL-1993 (Rel. 26, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Putative transporter yaaI.  
 GN YAAI OR B0007.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;

OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12;  
 RX MEDLINE-9233497; PubMed-1630901;  
 RA Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
 RA Isono K., Mizobuchi K., Nakata A.;  
 RT "Systematic sequencing of the Escherichia coli genome: analysis of  
 the 0-2.4 min region.";  
 RL Nucleic Acids Res. 20:3305-3308(1992).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE-97426617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:1453-1474(1997).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- SIMILARITY: BELONGS TO THE SODIUM:ALANINE SYMPORTER FAMILY  
 CC (SAP). STRONG. TO H. INFLUENZAE H10183.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: D10483; -; NOT\_ANNOTATED\_CDS.  
 DR EMBL: AE000111; AAC73118.1; -  
 DR Ecogene; EG1555; yaaI.  
 DR InterPro: IPR002293; Na\_al\_sympase\_1.  
 DR InterPro: IPR001463; Na\_al\_symp.  
 DR Pfam: PF01235; Na\_Ala\_symp. 1.  
 DR PRINTS: PR00175; NALASMPORT.  
 DR PROSITE: PS00873; NA\_ALANINE\_SYMPT. 1.  
 KW Hypothetical protein; Transmembrane; Inner membrane; Transport;  
 KW Symport; Complete proteome.  
 FT TRANSMEM 4 24  
 FT TRANSMEM 81 101  
 FT TRANSMEM 141 161  
 FT TRANSMEM 174 194  
 FT TRANSMEM 207 227  
 FT TRANSMEM 233 253  
 FT TRANSMEM 300 320  
 FT TRANSMEM 351 371  
 FT TRANSMEM 391 411  
 FT TRANSMEM 414 434  
 FT TRANSMEM 414 434  
 FT POTENTIAL.  
 FT SEQUENCE 476 AA; 51662 MW; 2F6EB2E12E126E63 CRC64;  
 Query Match 44.8%; Score 43; DB 1; Length 476;  
 Best Local Similarity 44.4%; Pred. No. 5.3;  
 Matches 8; Conservative 2; Mismatches 2; Indels 6; Gaps 1;  
 OY 1 DIN-----PAMVXXRGI 12  
 DB 120 DVNGQFRGGPAMYMARGL 137  
 RESULT 6  
 ID NRPI\_YEAST STANDARD; PRT; 719 AA.  
 AC P32770; Q12228;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)

```

DE  Apparagine-rich protein (ARP, protein).
GN  NRPI OR ARP OR YDL167C.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CX  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-AH22;
RX  MEDLINE=93247548; PubMed=8483449;
RA  Wehner E.P., Rao E., Brendel M.;
RT  "Molecular structure and genetic regulation of SRA, a gene
RT  responsible for resistance to formaldehyde in Saccharomyces
RT  cerevisiae, and characterization of its protein product.";
RL  MOL. Gen. Genet. 237:351-358(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-S288C;
RA  Pohl T.M.;
RL  Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.
CC  -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X68020; CAA48159.1; -
DR  EMBL; Z67750; CAA91579.1; -
DR  EMBL; Z74215; CAA98741.1; -
DR  PIR; S31139; S31139.
DR  HSSP; P04170; GRXN.
DR  SGD; S0002326; NRPI.
DR  InterPro; IPR000504; RRM.
DR  InterPro; IPR001876; Znf-RANBP.
DR  Pfam; PF00076; trm; 1.
DR  Pfam; PF00641; zf-RANBP; 2.
DR  SMART; SM00360; RRM; 1.
DR  SMART; SM00547; Znf-RB2; 2.
DR  PROSITE; PSS0102; RRM; 1.
DR  PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR  PROSITE; PS01358; ZF_RANBP2_1; 2.
DR  PROSITE; PSS0199; ZF_RANBP2_2; 2.
DR  Nucleic protein; zinc-finger; RNA-binding; Repeat.
FT  DOMAIN 226 322 RNA-BINDING (RRM).
FT  ZN_FING 355 384 RANBP2-TYPE 1.
FT  ZN_FING 581 610 RANBP2-TYPE 2.
FT  DOMAIN 490 564 ASN-RICH.
FT  CONFLICT 493 493 I -> N (IN REF. 1).
SQ  SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582669 CRC64;

Query Match 44.88; Score 43; DB 1; Length 719;
Best Local Similarity 50.0%; Pred. NO. 7.9;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

DE  (Exonuclease VII large subunit).
GN  XSEA OR SCK7.29C.
OS  Streptomyces coelicolor.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CX  Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX  NCBI_TaxID=1902;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-A3(2);
RA  Seeger K.J., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G.,
RA  Rajandream M.A.;
RL  Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE
CC  ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER
CC  INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' -
CC  or 3' - to 5'-direction to yield 5'-phosphomononucleotides.
CC  -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY
CC  SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AL391754; CAC05901.1; -
DR  InterPro; IPR003753; Exonuc_VII_L.
DR  InterPro; IPR002309; trna-synt_2.
DR  Pfam; PF02601; Exonuc_VII_L; 1.
DR  Pfam; PF01336; trna-anti; 1.
DR  Hydrolase; Nuclease; Exonuclease.
SQ  SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 43.88; Score 42; DB 1; Length 402;
Best Local Similarity 42.98; Pred. NO. 6.6;
Matches 9; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

```

```

DE  Apparagine-rich protein (ARP, protein).
GN  NRPI OR ARP OR YDL167C.
OS  Saccharomyces cerevisiae (Baker's yeast).
OC  Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
CX  Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX  NCBI_TaxID=4932;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-AH22;
RX  MEDLINE=93247548; PubMed=8483449;
RA  Wehner E.P., Rao E., Brendel M.;
RT  "Molecular structure and genetic regulation of SRA, a gene
RT  responsible for resistance to formaldehyde in Saccharomyces
RT  cerevisiae, and characterization of its protein product.";
RL  MOL. Gen. Genet. 237:351-358(1993).
RN  [2]
RP  SEQUENCE FROM N.A.
RC  STRAIN-S288C;
RA  Pohl T.M.;
RL  Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC  -1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.
CC  -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; X68020; CAA48159.1; -
DR  EMBL; Z67750; CAA91579.1; -
DR  EMBL; Z74215; CAA98741.1; -
DR  PIR; S31139; S31139.
DR  HSSP; P04170; GRXN.
DR  SGD; S0002326; NRPI.
DR  InterPro; IPR000504; RRM.
DR  InterPro; IPR001876; Znf-RANBP.
DR  Pfam; PF00076; trm; 1.
DR  Pfam; PF00641; zf-RANBP; 2.
DR  SMART; SM00360; RRM; 1.
DR  SMART; SM00547; Znf-RB2; 2.
DR  PROSITE; PSS0102; RRM; 1.
DR  PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR  PROSITE; PS01358; ZF_RANBP2_1; 2.
DR  PROSITE; PSS0199; ZF_RANBP2_2; 2.
DR  Nucleic protein; zinc-finger; RNA-binding; Repeat.
FT  DOMAIN 226 322 RNA-BINDING (RRM).
FT  ZN_FING 355 384 RANBP2-TYPE 1.
FT  ZN_FING 581 610 RANBP2-TYPE 2.
FT  DOMAIN 490 564 ASN-RICH.
FT  CONFLICT 493 493 I -> N (IN REF. 1).
SQ  SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582669 CRC64;

Query Match 44.88; Score 43; DB 1; Length 719;
Best Local Similarity 50.0%; Pred. NO. 7.9;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

```

```

DE  (Exonuclease VII large subunit).
GN  XSEA OR SCK7.29C.
OS  Streptomyces coelicolor.
OC  Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
CX  Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX  NCBI_TaxID=1902;
RN  [1]
RP  SEQUENCE FROM N.A.
RC  STRAIN-A3(2);
RA  Seeger K.J., Harris D., Cerdeno A.M., Parkhill J., Barrell B.G.,
RA  Rajandream M.A.;
RL  Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
CC  -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE
CC  ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER
CC  INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).
CC  -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' -
CC  or 3' - to 5'-direction to yield 5'-phosphomononucleotides.
CC  -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY
CC  SIMILARITY).
CC  -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC  -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.
CC  -----
CC  This SWISS-PROT entry is copyright. It is produced through a collaboration
CC  between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC  the European Bioinformatics Institute. There are no restrictions on its
CC  use by non-profit institutions as long as its content is in no way
CC  modified and this statement is not removed. Usage by and for commercial
CC  entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC  or send an email to license@isb-sib.ch).
CC  -----
DR  EMBL; AL391754; CAC05901.1; -
DR  InterPro; IPR003753; Exonuc_VII_L.
DR  InterPro; IPR002309; trna-synt_2.
DR  Pfam; PF02601; Exonuc_VII_L; 1.
DR  Pfam; PF01336; trna-anti; 1.
DR  Hydrolase; Nuclease; Exonuclease.
SQ  SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 43.88; Score 42; DB 1; Length 402;
Best Local Similarity 42.98; Pred. NO. 6.6;
Matches 9; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

```

RT Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL Science 273:1058-1073(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0576 AND TO S.POMBE MALARE  
PERMEASE (MAL1).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC at the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U67521; AAB98753.1; -  
DR TIGR: MJ0762; -  
KW Hypothetical protein; Transmembrane; Transport; Complete proteome.  
FT TRANSMEM 8 28 POTENTIAL.  
FT TRANSMEM 39 59 POTENTIAL.  
FT TRANSMEM 79 99 POTENTIAL.  
FT TRANSMEM 108 128 POTENTIAL.  
FT TRANSMEM 142 162 POTENTIAL.  
FT TRANSMEM 175 195 POTENTIAL.  
FT TRANSMEM 207 227 POTENTIAL.  
FT TRANSMEM 242 262 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 304 324 POTENTIAL.  
SQ SEQUENCE 342 AA; 39534 MW; 08FEFC3E2C4955D8 CRC64;  
  
Query Match 42.7%; Score 41; DB 1; Length 342;  
Best Local Similarity 53.3%; Pred. No. 8.5;  
Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;  
  
QY 2 INPAMYXXRGIRPVG 16  
Db 139 VNPGWY----IPVVG 149  
:||| | |||  
  
RESULT 9  
V576\_METJA STANDARD; PRT; 347 AA.  
ID V576\_METJA  
AC 057996;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein MJ0576.  
GN MJ0576.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., FitzGerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerecavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merrick J.M., Glodek A.,  
RA Scott J.L., Geoghegan N.S.M., Meldrum J.F., Furumman J.L., Nguyen D.,  
RA Usterbeck T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Clifton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klotz H.-P., Fraser C.M., Smith H.O., Moese C.R., Venter J.C.;  
RT \*Complete genome sequence of the methanogenic archaeon, Methanococcus  
jannaschii.";  
RL Science 273:1058-1073(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (potential).  
CC -1- SIMILARITY: STRONG, TO M.JANNASCHII MJ0762 AND TO S.POMBE MALARE  
PERMEASE (MAL1).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U67506; AAB98567.1; -  
DR TIGR: MJ0576; -  
KW Hypothetical protein; Transmembrane; Transport; Complete proteome.  
FT TRANSMEM 15 35 POTENTIAL.  
FT TRANSMEM 46 66 POTENTIAL.  
FT TRANSMEM 84 104 POTENTIAL.  
FT TRANSMEM 111 131 POTENTIAL.  
FT TRANSMEM 149 169 POTENTIAL.  
FT TRANSMEM 182 202 POTENTIAL.  
FT TRANSMEM 214 234 POTENTIAL.  
FT TRANSMEM 249 269 POTENTIAL.  
FT TRANSMEM 283 303 POTENTIAL.  
FT TRANSMEM 312 332 POTENTIAL.  
SQ SEQUENCE 347 AA; 39556 MW; 632EF7671A31DE183 CRC64;  
  
Query Match 42.7%; Score 41; DB 1; Length 347;  
Best Local Similarity 53.3%; Pred. No. 8.6;  
Matches 8; Conservative 1; Mismatches 2; Indels 4; Gaps 1;  
  
QY 2 INPAMYXXRGIRPVG 16  
Db 146 VNPGWY----IPVVG 156  
:||| | |||  
  
RESULT 10  
ENRN\_BPT7 STANDARD; PRT; 149 AA.  
ID ENRN\_BPT7  
AC P00641;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Endodeoxyribonuclease I (EC 3.1.21.2) (Endonuclease).  
GN 3.  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like phages.  
OX NCBI\_TaxID=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83241725; PubMed=6864790;  
RA Dunn J.J., Studier F.W.;  
RT \*Complete nucleotide sequence of bacteriophage T7 DNA and the  
RT locations of T7 genetic elements.";  
RL J. Mol. Biol. 166:477-535(1983).  
[2]  
RN SEQUENCE FROM N.A.  
RP MEDLINE=82078034; PubMed=7310871;  
RA Dunn J.J., Studier F.W.;  
RT \*Nucleotide sequence from the genetic left end of bacteriophage T7  
RT DNA to the beginning of gene 4.";  
RL J. Mol. Biol. 148:303-330(1981).  
CC -1- FUNCTION: ENDODEOXYRIBONUCLEASE I, WHICH IS EXPRESSED IN THE LATE  
CC STAGE, IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN  
CC OF HOST DNA. IN THE EARLY STAGE OF INFECTION, T7 DNA REPLICATES AS  
CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA  
CC LINEAR CONCATAMERS SEVERAL GENOMES IN LENGTH. THE GENE 3 PRODUCT  
CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE CONCATAMERS.  
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
CC phosphooligonucleotide end-products.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: V01127; CAA24345.1; -

DR EMBL: V01146; CAA24402.1; -

DR PIR: A00785; NEBP37.

DR PIR: S42301; S42301.

KW Hydrolase; Nuclease; Endonuclease.

SO SEQUENCE 149 AA; 17172 MW; D092AA28E3743BC1 CRC64;

Query Match 40.6%; Score 39; DB 1; Length 149;

Best Local Similarity 57.1%; Pred. No. 8.4;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 5 AMYXXRGIRPVGR 18

1 : ||| |||

Db 2 AGYAKGIRKYGAF 15

RESULT 11

SYE\_BACHD STANDARD; PRT; 485 AA.

ID SYE\_BACHD

AC Q9KGF; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE GlutamyL-trNA synthetase (EC 6.1.1.17) (Glutamate--trNA Ligase) (GluRS).

GN GLTX OR BH0109.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

CC Bacillus/staphylococcus group; Bacillus.

OX NCBI\_Taxid=86655;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;

RA MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S., Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + trNA(Glu) = AMP + diphosphate + L-glutamyl-trNA(Glu).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AP001507; BAB03828.1; -

DR Interpro: IPR000924; trNA-synt\_1c.

DR Interpro: IPR001412; trNA-synt\_1.

DR Pfam: PF00749; trNA-synt\_1c; 1.

DR PRINTS: PR00987; TRNASYNTHGLD.

DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.

KW Aminoacyl-trNA synthetase; Protein biosynthesis; Ligase; ATP-binding; Complete proteome.

FT SITE 11 21 "HIGH" REGION.

FT SITE 252 236 "KMSKS" REGION.

FT BINDING 255 255 ATP (BY SIMILARITY).

SO SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match 40.6%; Score 39; DB 1; Length 485;

Best Local Similarity 58.3%; Pred. No. 26;

Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 7 YXXRGIRPVGR 18

1 : ||| |||

Db 146 YEAKGIRPVGR 157

RESULT 12

SYV\_VIRBCH STANDARD; PRT; 953 AA.

ID SYV\_VIRBCH

AC Q9KPF; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Valyl-trNA synthetase (EC 6.1.1.9) (Valine--trNA Ligase) (ValRS).

GN VALS OR VC2503.

OS Vibrio cholerae.

OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.

OX NCBI\_Taxid=666;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=EL TOR N16961 / SEROTYPE O1;

RA MEDLINE=20406833; PubMed=10952301;

RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L., Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A., Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D., McDonallev M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P., McDonald L., Uppertack T., Fleischmann R.D., Nierman J.C., White O., Salzberg S.L., Smith R.O., Colwell R.R., Nekrasov J.J., Venter J.C., Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio cholerae.";

RL Nature 406:477-483(2000).

CC -1- CATALYTIC ACTIVITY: ATP + L-valine + trNA(Val) = AMP + diphosphate + L-valyl-trNA(Val).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: AE004320; AAF95645.1; -

DR HSP; P96142; IGAX.

DR TIGR: VC2503; -

DR Interpro: IPR002300; trNA-synt\_1a.

DR Interpro: IPR001412; trNA-synt\_1.

DR Interpro: IPR002303; trNA-synt\_val.

DR Pfam: PF00133; trNA-synt\_1; 1.

DR PRINTS: PR00986; TRNASYNTHVAL.

DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_1; 1.

KW Aminoacyl-trNA synthetase; Protein biosynthesis; Ligase; ATP-binding; Complete proteome.

FT SITE 42 52 "HIGH" REGION.

FT SITE 554 558 "KMSKS" REGION.

FT BINDING 557 557 ATP (BY SIMILARITY).

SO SEQUENCE 953 AA; 108170 MW; D93471A33CF4F69C CRC64;

Query Match 40.6%; Score 39; DB 1; Length 953;

Best Local Similarity 57.1%; Pred. No. 51;

Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 4 PAWYXXRGIRPVGR 17

|||||

Db 436 PAWYDEGQNVFVR 449

RESULT 13  
ID DPOL\_THERFL STANDARD: PRT: 831 AA.  
AC P3013:  
DR 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DE 15-DEC-1998 (Rel. 37, Last annotation update)  
GN POLA OR POL.  
OS Thermus aquaticus (subsp. flavus).  
OC Bacteria; Thermus/Deinococcus group; Thermus group; Thermus.  
OX NCBI\_TaxID=274;  
RN NCBI\_TaxID=274;  
RP SEQUENCE FROM N.A.  
RX STRAIN-ACM B-1257;  
RA MEDLINE=93087201; PubMed=1454544;  
RT Akhmetzhanov A.A., Vakhitov V.A.;  
RT "Molecular cloning and nucleotide sequence of the DNA polymerase gene  
from Thermus flavus.";  
RL Nucleic Acids Res. 20:5839-5839(1992).  
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate  
+ (DNA)(N).  
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: X66105; CAA46900.1; -  
DR PIR: S24929; S24929.  
DR PIR: S26675; S26675.  
DR HSSP: P19821; ITAO.  
DR InterPro: IPR002421; 5\_3-exonuclease.  
DR InterPro: IPR002298; DNA\_pol.  
DR InterPro: IPR001098; DNA\_pol\_A.  
DR InterPro: IPR000513; Exo\_N.I.  
DR InterPro: IPR003583; HHH\_1.  
DR InterPro: IPR003584; HHH\_2.  
DR InterPro: IPR001532; XPG\_1.  
DR Pfam: PF01367; 5\_3-exonuclease; 1.  
DR Pfam: PF02739; 5\_3-exonuc N; 1.  
DR Pfam: PF00476; DNA\_pol\_A; 1.  
DR PRINTS: PR00868; DNAPOL.  
DR SMART: SM00475; 53EXOC; 1.  
DR SMART: SM00278; HhH1; 1.  
DR SMART: SM00279; HhH2; 1.  
DR SMART: SM00483; POLAC; 1.  
DR SMART: SM00485; XPGN; 1.  
DR PROSITE: PS00447; DNA-POLYMERASE\_A; 1.  
DR Transferrase; DNA-directed DNA polymerase; DNA replication; DNA repair;  
KW DNA-binding.  
KW DOMAIN  
FT DOMAIN 409 831 POLYMERASE (BY SIMILARITY).  
FT SEQUENCE 831 AA; 93783 MW; 96F93CEFA3CA536D CRC64;  
OY 2 INPAW-YXXRGTRP 14  
Db 162 ITPAWLYEKYGRP 175  
Query Match 40.1%; Score 38.5; DB 1; Length 831;  
Best Local Similarity 57.1%; Pred. No. 55;  
Matches 8; Conservative 1; Mismatches 4; Indels 1; Gaps 1;

DR 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DE 16-OCT-2001 (Rel. 40, Last annotation update)  
GN Neural-cadherin precursor (N-cadherin) (Cadherin-2) (Fragment).  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN NCBI\_TaxID=9913;  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90360979; PubMed=2390969;  
RA Liao C.W., Cannon C., Power M.D., Kiboneka P.K., Rubin L.L.;  
RT "Identification and cloning of two species of cadherins in bovine  
endothelial cells.";  
RL EMBO J. 9:2701-2708(1990).  
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
CC NEURONAL RECOGNITION MECHANISM.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
CC -----  
DR EMBL: X5615; CAA37677.1; -  
DR PIR: S11693; IJBOCN.  
DR HSSP: P15116; INCI.  
DR InterPro: IPR002126; Cadherin.  
DR InterPro: IPR002333; Cadherin\_C-term.  
DR Pfam: PF00028; cadherin; 5.  
DR Pfam: PF01049; Cadherin\_C-term; 1.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA; 5.  
DR PROSITE: PS00232; CADHERIN\_1; 3.  
DR PROSITE: PS00268; CADHERIN\_2; 5.  
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.  
KW CELL ADHESION; GLYCOPROTEIN; TRANSMEMBRANE; CALCIUM-BINDING; REPEAT.  
FT PROPEP 1 130  
FT CHAIN 131 877  
FT DOMAIN 131 695  
FT TRANSMEM 696 717  
FT DOMAIN 718 877  
FT DOMAIN 131 238  
FT DOMAIN 239 363  
FT DOMAIN 364 468  
FT DOMAIN 469 574  
FT DOMAIN 575 685  
FT DOMAIN 685 849  
FT CARBOHYD 161 161  
FT CARBOHYD 244 244  
FT CARBOHYD 296 296  
FT CARBOHYD 373 373  
FT CARBOHYD 543 543  
FT CARBOHYD 622 622  
FT CARBOHYD 663 663  
FT SEQUENCE 877 AA; 96845 MW; 441B829DB871A249 CRC64;  
OY 4 PAW---YXXRGTRPGRF 18  
Query Match 40.1%; Score 38.5; DB 1; Length 877;  
Best Local Similarity 50.0%; Pred. No. 58;  
Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

Db 385 PAMNAVIRISGDPGRF 402

RESULT 15

ID CAD2\_HUMAN STANDARD; PRT; 906 AA.

AC P19022; 014923; 01-NOV-1990 (Rel. 16, Created)

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).

GN CDH2 OR CDHN OR NCAD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID:9606;

RN (1)

RP SEQUENCE FROM N.A.

RA Reid R.A., Hemperly J.J.;

RT "Human N-cadherin: nucleotide and deduced amino acid sequence.";

RL Nucleic Acids Res. 18:5896-5896(1990).

RN (2)

RP REVISIONS TO 341: 699 AND 705.

RA Reid R.A.;

RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.

RN (3)

RP SEQUENCE FROM N.A.

RA MEDLINE-92363956; PubMed-1500442;

RA Salomon D., Avalon O., Patel-King R., Hynes R.O., Gelber B.;

RT "Extrajunctional distribution of N-cadherin in cultured human endothelial cells.";

RL J. Cell Sci. 102:7-17(1992).

RN (4)

RP SEQUENCE OF 160-906 FROM N.A.

RA MEDLINE-90347462; PubMed-2384753;

RA Walsh F.S., Barton C.H., Putt W., Moore S.E., Kelsell D.;

RA Spurr N., Goodfellow P.N.;

RT "N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin gene.";

RL J. Neurochem. 55:805-812(1990).

RN (5)

RP SEQUENCE OF 1-20 FROM N.A.

RA MEDLINE-95048366; PubMed-7959764;

RA Wallis J.A., Fox M., Walsh F.S.;

RT "Structure of the human N-cadherin gene: YAC analysis and fine chromosomal mapping to 18q11.2.";

RL Genomics 22:172-179(1994).

CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS. THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN NEURONAL RECOGNITION MECHANISM.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

DR EMBL: X57548; CAA40773.1; -

DR EMBL: X54315; CAA38213.1; -

DR EMBL: S43303; AAB23854.1; -

DR EMBL: M34064; AAA03236.1; -

DR EMBL: Z27420; CAA81799.1; -

DR PIR: A38870; IJHUCN.

DR HSSP: P15116; INCI.

DR MIM: 114020; -

DR InterPro: IPR002126; Cadherin.

DR InterPro: IPR002126; Cadherin\_C-term.

DR Pfam: PF00028; cadherin\_5.

DR Pfam: PF01049; Cadherin\_C-term; 1.

DR PRINTS: PR00205; CADHERIN.

DR SMART: SM00112; CA; 5.

DR PROSITE: PS00232; CADHERIN\_1; 3.

DR PROSITE: PS02068; CADHERIN\_2; 5.

DR Cell adhesion; glycoprotein; Transmembrane; Calcium-binding; Repeat; signal.

FW SIGNAL. 1 23

FT PROPEP 24 159

FT CHAIN 160 906

FT DOMAIN 160 724

FT TRANSMEM 725 746

FT DOMAIN 747 906

FT DOMAIN 160 267

FT DOMAIN 268 382

FT DOMAIN 383 497

FT DOMAIN 498 603

FT DOMAIN 604 714

FT DOMAIN 863 878

FT CARBOHYD 190 190

FT CARBOHYD 273 273

FT CARBOHYD 325 325

FT CARBOHYD 402 402

FT CARBOHYD 572 572

FT CARBOHYD 651 651

FT CARBOHYD 692 692

FT CONFLICT 12 12

FT CONFLICT 16 16

FT CONFLICT 196 196

FT CONFLICT 212 212

FT CONFLICT 357 357

FT CONFLICT 867 867

SO SEQUENCE 906 AA; 99851 MW; 72DDC7BB857C7AFC CRC64;

Query Match 40.18; Score 38.5; DB 1; Length 906;

Best Local Similarity 50.0%; Pred. No. 59;

Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 4 PAM---YXXRGIRPVGR 18

Db 414 PAMNAVIRISGDPGRF 431

Search completed: September 13, 2002, 09:30:45

Job time: 1135 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:19 ; Search time 311.85 Seconds

(without alignments)  
10.540 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_3\_21

Perfect score: 96

Sequence: 1 DPNPAMYXXRGIRPVGRFX 19

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP\_ARCHAEA:\*  
2: SP\_BACTERIA:\*  
3: SP\_FUNGI:\*  
4: SP\_HUMAN:\*  
5: SP\_INVERTEBRATE:\*  
6: SP\_MAMMAL:\*  
7: SP\_MHC:\*  
8: SP\_ORGANELLE:\*  
9: SP\_PHAGE:\*  
10: SP\_PLANT:\*  
11: SP\_RODENT:\*  
12: SP\_VIRUS:\*  
13: SP\_VERTEBRATE:\*  
14: SP\_UNCLASSIFIED:\*  
15: SP\_VIRUS:\*  
16: SP\_BACTERIAP:\*  
17: SP\_ARCHAEP:\*

Pred. NO. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query	Match Length	DB ID	Description
1	74	77.1	117	13	Q9W624
2	46.5	48.4	420	2	Q93127
3	46	47.9	333	16	Q9PH76
4	46	47.9	540	10	Q9L620
5	46	47.9	790	10	Q9M371
6	44	45.8	419	4	Q9Y276
7	43	44.8	220	16	Q91022
8	43	44.8	418	11	Q9C2B5
9	42.5	44.3	360	5	Q19879
10	42	43.8	416	16	Q99Z89
11	42	43.8	433	16	Q9A382
12	41	42.7	97	2	Q33440
13	41	42.7	226	16	Q9A359
14	41	42.7	280	4	Q96925
15	41	42.7	328	4	Q9NVR5
16	41	42.7	338	5	Q18729

17	41	42.7	545	16	Q9A7W7	Q9A7W7 caulobacter
18	40	41.7	76	5	Q9V8L0	Q9V8L0 drosophila
19	40	41.7	153	9	Q9L133	Q9L133 bacteriophage
20	40	41.7	184	2	Q96838	Q96838 streptomyces
21	40	41.7	244	10	Q9S1D9	Q9S1D9 arabidopsis
22	40	41.7	284	17	Q50128	Q50128 pyrococcus
23	40	41.7	289	10	Q9F0S5	Q9F0S5 oryza sativa
24	40	41.7	324	2	Q9Z554	Q9Z554 streptomyces
25	40	41.7	390	16	Q9PH18	Q9PH18 xylella fastidiosa
26	40	41.7	430	17	Q27142	Q27142 methanococcus
27	40	41.7	446	10	Q945R7	Q945R7 oryza sativa
28	40	41.7	462	10	Q9Z249	Q9Z249 arabidopsis
29	40	41.7	486	16	Q92DM2	Q92DM2 listeria in
30	40	41.7	562	5	Q9VNG4	Q9VNG4 drosophila
31	39.5	41.1	779	3	Q9P5J9	Q9P5J9 neurospora
32	39	40.6	105	16	Q9PNE9	Q9PNE9 campylobacter
33	39	40.6	108	11	Q9D1V4	Q9D1V4 mus musculus
34	39	40.6	179	2	Q9K3Q7	Q9K3Q7 streptomyces
35	39	40.6	250	16	Q912F2	Q912F2 pseudomonas
36	39	40.6	256	16	Q92VE9	Q92VE9 rhizobium m
37	39	40.6	283	2	Q932V3	Q932V3 campylobacter
38	39	40.6	327	10	Q9L0I4	Q9L0I4 arabidopsis
39	39	40.6	333	4	Q96SD4	Q96SD4 homo sapien
40	39	40.6	340	5	Q18731	Q18731 caenorhabditis
41	39	40.6	370	11	Q9D3V7	Q9D3V7 mus musculus
42	39	40.6	425	16	Q986U6	Q986U6 rhizobium l
43	39	40.6	555	3	Q00050	Q00050 aspergillus
44	39	40.6	938	10	Q9ZVE3	Q9ZVE3 arabidopsis
45	39	40.6	1488	5	Q20294	Q20294 caenorhabditis

## ALIGNMENTS

RESULT 1					
Q9W624	1	PRELIMINARY;	PRT;	117 AA.	
AC Q9W624;					
DT 01-NOV-1999 (TREMBLrel. 12, Created)					
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)					
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					
DE C-RR AMIDE.					
OS Carassius auratus (Goldfish).					
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;					
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;					
OC Cypriniformes; Cyprinidae; Carassius.					
OX NCBI_TaxID=7957;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE=BRAIN;					
RA Satake H., Minakata H., Fujimoto M.;					
RT "Carassius Rhamde (C-RR amide)".					
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.					
DR EMBL; AB020024; BAA/6662.1; -.					
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;					

Query Match	77.1%	Score 74:	DB 13:	Length 117:
Best Local Similarity	61.1%	Pred. No. 2.8e-05:		
Matches 11: Conservative		4: Mismatches	3: Indels	0: Gaps
QY 1 DPNPAMYXXRGIRPVGRFX 18				
DB 58 EIDPFMYVGRGVPRGRFX 75				
RESULT 2				
Q93L27	2	PRELIMINARY;	PRT;	420 AA.
AC Q93L27;				
DT 01-DEC-2001 (TREMBLrel. 19, Created)				
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)				
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)				

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
 GN AUR2B.  
 OS Streptomyces aureofaciens.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1894;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCM3239;  
 RA Kormanec J., Bistakova J., Novakova R., Homeroova D., Rezuchova B.;  
 RT "Cloning and characterization of a new polyketide gene cluster in  
 RT Streptomyces aureofaciens CCM3239".  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY033994; AAK61719.1;  
 SQ SEQUENCE 420 AA; 43011 MW; 3C27E2B8C2DEA CRC64;

Query Match 48.4%; Score 46.5; DB 2; Length 420;  
 Best Local Similarity 52.6%; Pred. No. 6.5;  
 Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

OY 5 AMYXX-----RGIRPVGRF 18  
 |||  
 Db 37 AMMAAVLKGESGIRPVGRF 55

RESULT 3  
 O9P176 PRELIMINARY; PRT; 333 AA.  
 AC O9PH76;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HYDROXYBENZONATE OCTAPRENYLTRANSFERASE.  
 GN XF0068.  
 OS Xylella fastidiosa.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2171;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RC MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Briones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
 RA Colaço N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
 RA Falcinani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Frega J.S., Franca S.C., Franco M.C., Frohne M., Furian L.R.,  
 RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuranse E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Merck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terezzi M.F., Truffi D., Tsesi S.M., Tsubako M.H.,  
 RA Valada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa".  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003860; AAF82881.1;  
 DR InterPro: IPR000537; UblA.

DR Pfam; PF01040; UblA; 1.  
 DR PROSITE; PS00943; UblA; UNKNOWN\_1.  
 KW Complete Proteome.  
 SQ SEQUENCE 333 AA; 37931 MW; ECF3F4716C962B95 CRC64;

Query Match 47.9%; Score 46; DB 16; Length 333;  
 Best Local Similarity 53.3%; Pred. No. 6.2;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 INPAMXXRGIRPVG 16  
 ::|||  
 Db 54 LDPTWKLARGDRPVG 68

RESULT 4  
 O9LGZ0 PRELIMINARY; PRT; 540 AA.  
 ID O9LGZ0;  
 AC O9LGZ0;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE EST\_A070209(R3722) CORRESPONDS TO A REGION OF THE PREDICTED  
 DE GENE.  
 OS Oryza sativa (Rice).  
 OS Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoidae; Oryzeae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponebare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:p0702f03".  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER  
 CC PROTEINS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE = AMP +  
 CC PYROPHOSPHATE + PROTEIN N-UBIQUITYLYSINE.  
 CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.  
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-  
 CC THIOLESTER FORMATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
 DR EMBL: AP002481; BAA96583.1;  
 DR HSSP: P06104; IAY2.  
 DR InterPro: IPR000608; UBO\_conjugat.  
 DR Pfam: PF00179; UO\_con; 1.  
 DR SMART: SM00212; UBCC; 1.  
 DR PROSITE: PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
 DR LIGase; Ubiqultin conjugation.  
 SQ SEQUENCE 540 AA; 60487 MW; 5DE1PF4EBB75A86E CRC64;

Query Match 47.9%; Score 46; DB 10; Length 540;  
 Best Local Similarity 50.0%; Pred. No. 10;  
 Matches 7; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 1 DINPAMXXRGIRP 14  
 |::|||  
 Db 69 DLGVAMWVRGLRP 82

RESULT 5  
 O9M371 PRELIMINARY; PRT; 790 AA.  
 ID O9M371;  
 AC O9M371;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HYPOTHETICAL 87.4 KDA PROTEIN.  
 GN F15616.60.  
 OS Arabidopsis thaliana (Mouse-ear cross).

```

OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,
RA Mayer K.F.X., Queller F., Salanoubat M.;
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.
DR EMBL; AL132959; CAB71097.1; -.
KW Hypothetical protein.
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

```

RT tubulopathy and liver failure.";  
 RN Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases  
 RL [5]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=RHABDOMYOSARCOMA;  
 RA Strausberg R.;  
 RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases  
 RN [6]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=MUSCLE, RHABDOMYOSARCOMA;  
 RA Strausberg R.;  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases  
 DR EMBL; AF026849; AAD08638.1; -.  
 DR EMBL; AF038195; AAB97365.1; -.  
 DR EMBL; AF346835; AAK29417.1; -.  
 DR EMBL; BC000416; AAH00416.1; -.  
 DR EMBL; BC007500; AAH07500.1; -.  
 DR InterPro: IPR003593; AAA.  
 DR InterPro: IPR003959; AAA\_subfam.  
 DR Pfam: PF00004; AAA; 1.  
 DR SMART; SM00382; AAA; 1.  
 DR ATP-binding; Hypothetical protein.  
 SW SEQUENCE 419 AA: 4753 MW: 7F0F98BA62E2CBB8 CRC64;

RESULT	6			
09Y276		PRELIMINARY;	PRT;	419 AA.
AC	09Y276			
AC	09Y276;			
DT	01-NOV-1999 (TREMBLrel. 12, created)			
DT	01-NOV-1989 (TREMBLrel. 12, last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 15, last annotation update)			
DE	H-BCS1 (BCS1 (YEAST HOMOLOG) -like).			
GN	BCS1 OR BCS1L.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=BRAIN;			
RC	MEDLINE=96207350; PubMed=9878253;			
RX	peruzzella V., Tifantl V., Fernandez P., Ianna P., Carozzo R.,			
RA	zeytani M.;			
RA	"identification and characterization of human cDNAs specific to BCS1			
RT	Pe112, SC01, COX15, and COX11, five genes involved in the formation			
RT	and function of the mitochondrial respiratory chain.";			
RL	Genomics 54:494-504(1998).			
RN	[2]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=BRAIN;			
RC	MEDLINE=96207227; PubMed=8619474;			
RX	Anderson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;			
RT	"A 'double adaptor' method for improved shotgun library			
RT	construction.";			
RL	Anal. Biochem. 236:107-113(1996).			
RN	[3]			
RN	SEQUENCE FROM N.A.			
RP	TISSUE=BRAIN;			
RC	MEDLINE=97264341; PubMed=9110174;			
RX	Yu W., Anderson B., Worley R.C., Muzny D.M., Ding Y., Liu W.,			
RA	Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;			
RA	"large-scale concatenation cDNA sequencing.";			
RL	Genome Res. 7:353-358(1997).			
RN	[4]			
RN	SEQUENCE FROM N.A.			
RP	de Lonlay P., Vainlot I., Barrientos A., Gorbatyuk M., Tzagoloff A.,			
RA	Benayoun E., Chretien D., Kadhom N., Lombes A., Ogier de Baulny H.,			
RA	Niaudet P., Munnich A., Rustin P., Rotig A.;			
RT	"Mutations in bcs1, a mitochondrial respiratory chain assembly gene,			
RT	are responsible for the complex III deficiency of patients with			

Query Match	45.8%	Score 44:	DB 4:	Length 419:
Best Local Similarity	70.0%	Pred. No. 18:		
Matches 7:	Conservative	0:	Mismatches 3:	Indels 0:
				Gaps 0:
OY	3	NPAMYXXRGI	12	
Db	211	NPKWTDRGI	220	
RESULT	7			
ID	091022	PRELIMINARY:	PRT:	220 AA.
AC	091022:			
DT	01-MAR-2001 (TREMBLrel. 16, Created)			
DT	01-MAR-2001 (TREMBLrel. 16, Last sequence update)			
DE	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DI	PROBABLE GLUTATHIONE S-TRANSFERASE.			
GN	PA2821.			
OS	Pseudomonas aeruginosa.			
OC	Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;			
OC	Pseudomonas.			
OX	NCBI_TaxID=287;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN-ATCC 15692 / PA01;			
RX	MEDLINE=20437337; PubMed=10984043;			
RA	Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,			
RA	Hickey M.J., Brinkman F.S.L., Huffnagle W.O., Kowalik D.J., Lagrou M.,			
RA	Gardner R.L., Goltzy L., Tolentino E., Westbrock-Medman S., Yuan Y.,			
RA	Brody L.L., Coulter S.N., Folger K.R., Kas A., Lardig K., Lim R.M.,			
RA	Smith K.A., Spencer D.H., Wong G.K.-S., Wong Z., Paulsen I.T.,			
RA	Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;			
RT	"Complete genome sequence of Pseudomonas aeruginosa PA01, an			
RT	opportunistic pathogen."			
RL	Nature 406:959-964(2000).			
DR	EMBL: AE004709; AAC06209.1; -			
DR	InterPro: IPR004046; GSF_C.			
DR	InterPro: IPR004045; GSF_N.			
DR	Transferrase; Complete proteome.			
SO	SEQUENCE 220 AA; 24716 MW; 6596183EA6CAA050 CRC64;			
Query Match	44.8%	Score 43:	DB 16:	Length 220;
Best Local Similarity	64.3%	Pred. No. 13;		
Matches 9:	Conservative	1:	Mismatches 2:	Indels 2:
				Gaps 1:
OY	4	PAWYXXRGI	PVGR 17	

Db 39 PAMV--REISPLGR 50

RESULT 8

09CZP5 PRELIMINARY: PRT: 418 AA.

AC 09CZP5

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE 9130022019RIK PROTEIN.

GN 9130022019RIK

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

OX NCBI\_TaxID=10090;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C57BL/6J; TISSUE=EMBryo;

RX MEDLINE=21085660; PubMed=11217851;

RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,

RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,

RA Atzawa K., Itawa M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,

RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,

RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,

RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,

RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,

RA Schirral L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,

RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,

RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,

RA Brownstein M.J., Bull C., Fletcher C., Fujita M., Gariboldi M.,

RA Guelincich S., Hill D., Hofmann M., Hume D.A., Kamliya M., Lee N.H.,

RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,

RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,

RA Norokoshi H., Sato K., Schenbach C., Seta T., Shibata Y., Storch K.-F.,

RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilmink L.,

RA Wystraw-Boris A., Yoshida K., Hasegawa Y., Kawai H., Kontecki S.,

RA Hayashizaki Y.;

RT "Functional annotation of a full-length mouse cDNA collection.";

RL Nature 409:685-690(2001).

DR EMBL: AK012324; BAB28162.1; -

DR MGD: MGI:1914071; 9130022019RIK.

DR InterPro: IPR003593; AAA.

DR InterPro: IPR003593; AAA-subfam.

DR Pfam: PF00004; AAA; 1.

DR SMART: SM00382; AAA; 1.

KW ATP-binding.

SO SEQUENCE 418 AA; 47406 MW; 94905BA9B097F0DE CRC64;

Query Match 44.8%; Score 43; DB 11; Length 418;

Best Local Similarity 70.0%; Pred. No. 26;

Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 NPAMYXXRGI 12

Db 211 NPKWYIDRGI 220

RESULT 9

019879 PRELIMINARY: PRT: 360 AA.

ID 019879

AC 019879

DT 01-NOV-1996 (TREMBLrel. 01, Created)

DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE F2BD1.8 PROTEIN.

GN F2BD1.8.

OS Caenorhabditis elegans.

OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;

OC Rhabditidae; Pelodierinae; Caenorhabditis.

OX NCBI\_TaxID=6239;

RN [1]

RP SEQUENCE FROM N.A.

RA Baynes C.;

RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.

RN [2]

RP SEQUENCE FROM N.A.

RX MEDLINE=99069613; PubMed=9851916;

RA none;

RT "Genome sequence of the nematode C.elegans: A platform for

RT investigating biology.";

RL Science 282:2012-2018(1996).

DR EMBL: Z70684; CAA94603.2; -

SO SEQUENCE 360 AA; 39234 MW; A5BBE127E7B3A91C CRC64;

Query Match 44.3%; Score 42.5; DB 5; Length 360;

Best Local Similarity 37.5%; Pred. No. 27;

Matches 9; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

OY 2 INPAMYXXR-----GIRPVG 16

Db 53 INPSWFDWRVSVSIDGHIGWPIG 76

RESULT 10

0992A9 PRELIMINARY: PRT: 416 AA.

AC 0992A9

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PUTATIVE EXTRAMEMBRANAL PROTEIN.

GN DLYD OR SPY1309.

OS Streptococcus pyogenes.

OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;

CC Streptococcus.

OX NCBI\_TaxID=1314;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;

RX MEDLINE=21192684; PubMed=11296296;

RA Ferretti J.J., Moshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,

RA Primeaux C., Szate S., Suvorov A.N., Kenton S., Lal H.S., Lin S.P.,

RA Yuan Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,

RA Qian Y., Clifton S.W., Roe B.A., McLaughlin R.;

RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";

RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).

DR EMBL: AE006570; AAK34153.1; -

KW Complete proteome.

SO SEQUENCE 416 AA; 48068 MW; EACAA2203F7E519F CRC64;

Query Match 43.8%; Score 42; DB 16; Length 416;

Best Local Similarity 38.5%; Pred. No. 39;

Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 2 INPAMYXXRGI 14

Db 129 VSPQWFTAGI 141

RESULT 11

09A382 PRELIMINARY: PRT: 433 AA.

ID 09A382

AC 09A382

DT 01-JUN-2001 (TREMBLrel. 17, Created)

DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE PEPTIDOGLYCAN-BINDING PROTEIN, PUTATIVE.

GN CC3322.

OS Caulobacter crescentus.

OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;

CC Caulobacter.

OX NCBI\_TaxID=69394;

[1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN-ATCC 19089 / CB15;  
 RC MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Ueberback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005994; AAK25284.1; -.  
 DR HSSP; P41052; 1LTM.  
 DR TIGR; CC3332; -.  
 DR InterPro: IPR002477; PG-binding.  
 DR Pfam: PF01471; PG-binding\_1; 1.  
 KM Complete proteome.  
 SQ SEQUENCE 433 AA; 46169 MW; F7DD1BB9FACA10A7 CRC64;

Query Match 43.8%; Score 42; DB 16; Length 433;  
 Best Local Similarity 60.0%; Pred. No. 41;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 4 PAMWXXRGIR 13  
 DB 285 PAMWEARGVR 294

RESULT 12  
 ID 033440 PRELIMINARY; PRT; 97 AA.  
 AC 033440:  
 DT 01-JAN-1998 (TREMblrel. 05, Created)  
 DT 01-JAN-1998 (TREMblrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE NADH-UBIQUINONE OXIDOREDUCTASE (FRAGMENT).  
 GN NUO4.  
 OS Pseudomonas fluorescens.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
 OC Pseudomonas.  
 OX NCBI\_TaxID=294;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-WCS365;  
 RX MEDLINE-98340543; PubMed-9675892;  
 RA Dekkers L.C., van der Bij A.J., Mulders H.M., Phoelich C.C.,  
 RA Wentwood A.R., Glandorf D.C.M., Wijffelman C.A., Lugtenberg B.J.J.;  
 RT "Role of the O-antigen of lipopolysaccharide, and possible roles of  
 RT growth rate and NADH:ubiquinone oxidoreductase (nuo) in competitive  
 RT tomo to root-tip colonization by *Pseudomonas fluorescens* WCS365.",  
 RL Mol. Plant Microbe Interact. 11:763-771(1998).  
 DR EMBL; Y14569; GAU74900.1; -.  
 DR InterPro: IPR001135; Complex1\_49kd.  
 DR Pfam: PF00346; complex1\_49kd; 1.  
 KW Ubiquinone; NAD.  
 FT NON\_TER  
 FT NON\_TER  
 SQ SEQUENCE 97 AA; 11450 MW; 656E8C662A3939D9 CRC64;

Query Match 42.7%; Score 41; DB 2; Length 97;  
 Best Local Similarity 46.2%; Pred. No. 12;  
 Matches 6; Conservative 3; Mismatches 4; Indels 0; Gaps 0;

OY 2 INPAMWXXRGIRP 14  
 DB 54 LHPAMWYRGVCP 66

RESULT 13

09A359  
 ID 09A359 PRELIMINARY; PRT; 226 AA.  
 AC 09A359:  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE HYPOTHETICAL PROTEIN CC3347.  
 GN CC3347.  
 OS *Caulobacter crescentus*.  
 OC Bacteria; Proteobacteria; alpha subdivision; *Caulobacter* group;  
 OC *Caulobacter*.  
 OX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE-21173698; PubMed-11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 RA Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
 RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Ueberback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of *Caulobacter crescentus*.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005996; AAK25309.1; -.  
 DR TIGR; CC3347; -.  
 KM Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 226 AA; 25010 MW; CD9C0AACC3B26CD CRC64;

Query Match 42.7%; Score 41; DB 16; Length 226;  
 Best Local Similarity 41.2%; Pred. No. 30;  
 Matches 7; Conservative 3; Mismatches 7; Indels 0; Gaps 0;

OY 2 INPAMWXXRGIRPVGRF 18  
 DB 125 VNPDWMSGRALRDVQF 141

RESULT 14  
 ID 096925 PRELIMINARY; PRT; 280 AA.  
 AC 096925:  
 DT 01-DEC-2001 (TREMblrel. 19, Created)  
 DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE HYPOTHETICAL 31.4 KDA PROTEIN (SIMILAR TO HYPOTHETICAL PROTEIN  
 DE FLJ10563).  
 OS *Homo sapiens* (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BONE MARROW, AND ACUTE MYELOCYTIC LEUKEMIA;  
 RA Strausberg R.;  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-COLON ADENOCARCINOMA;  
 RA Strausberg R.;  
 RL Submitted (JUL-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; BC013322; AAH13322.1; -.  
 DR EMBL; BC011400; AAH11400.1; -.  
 KW Hypothetical protein.  
 SQ SEQUENCE 280 AA; 31429 MW; 4A22CCA331F50DA6 CRC64;

Query Match 42.7%; Score 41; DB 4; Length 280;  
 Best Local Similarity 60.0%; Pred. No. 38;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 DINPAWYXXR 10  
 | : | | | | |  
 Db 42 DNNPLMYKLR 51

RESULT 15

09NVR5 PRELIMINARY; PRT; 328 AA.  
 ID 09NVR5;  
 AC 09NVR5;  
 DT 01-OCT-2000 (TREMBlurel. 15, Created)  
 DT 01-OCT-2000 (TREMBlurel. 15, last sequence update)  
 DT 01-JUN-2001 (TREMBlurel. 17, last annotation update)  
 DE CDNA FLJ10563 FIS, CLONE NTZRP2002769.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Isogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Takahashi-Fujii A., Hara H.,  
 RA Tanase T., Nomura Y., Togiya S., Komai F., Hara R., Takeuchi K.,  
 RA Arita M., Nabekura T., Ishii S., Kawai Y., Saito K., Yamamoto J.,  
 RA Wakamatsu A., Nakamura Y., Nagahari K., Masuho Y., Oshima A.,  
 RT "NEDO human cDNA sequencing project."  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AK001425; BAA91684.1; -  
 SQ SEQUENCE 328 AA; 36943 MW; 87875CE2A89AF663 CRC64;

Query Match 42.78; Score 41; DB 4; Length 328;  
 Best Local Similarity 60.08; Pred. No. 45;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

Oy 1 DINPAWYXXR 10  
 | : | | | | |  
 Db 42 DNNPLMYKLR 51

Search completed: September 13, 2002, 09:29:19  
 Job time: 1064 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds  
(without alignments)  
5,002 Million cell updates/sec

Title: us-09-446-543a-73\_COPY\_4\_21  
Perfect score: 90  
Sequence: 1 INPAMYXXRCIRPYGRPX 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_032802:\*

- 1: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSL/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	96.7	20	AAW31394	Human type G prote
2	87	96.7	20	AAW97235	Human type ligand
3	87	96.7	20	AAW10365	Human oxytocin sec
4	87	96.7	20	AAV49294	1992 ligand peptid
5	87	96.7	20	AAW62534	Human CRH releasin
6	87	96.7	20	AAW90992	Human type G prote
7	87	96.7	21	AAW31395	Human oxytocin sec
8	87	96.7	21	AAW10366	Human CRH releasin
9	87	96.7	21	AAW62535	Human type G prote
10	87	96.7	22	AAW31396	Human oxytocin sec
11	87	96.7	22	AAW10367	Human oxytocin sec

12	87	96.7	22	AAW62536	Human CRH releasin
13	87	96.7	31	AAW31391	Human type G prote
14	87	96.7	31	AAW97235	Human type ligand
15	87	96.7	31	AAW87615	Human 1992 ligand.
16	87	96.7	31	AAW10362	Human oxytocin sec
17	87	96.7	31	AAV49291	1992 ligand peptid
18	87	96.7	31	AAW62531	Human CRH releasin
19	87	96.7	31	AAW90991	Human CRH releasin
20	87	96.7	31	AAW90995	Human type G prote
21	87	96.7	32	AAW31392	Human oxytocin sec
22	87	96.7	32	AAW10363	Human CRH releasin
23	87	96.7	32	AAW62532	Human type G prote
24	87	96.7	33	AAW10364	Human type G prote
25	87	96.7	33	AAW62533	Human CRH releasin
26	87	96.7	33	AAW97232	Human oxytocin sec
27	87	96.7	37	AAW31390	Human type G prote
28	87	96.7	87	AAW97226	Human type ligand
29	87	96.7	87	AAW10361	Human CRH releasin
30	87	96.7	87	AAW62530	Human CRH releasin
31	86	95.6	20	AAW31387	Rat type G protein
32	86	95.6	20	AAW31374	Bovine G protein-c
33	86	95.6	20	AAW97232	Bovine pituitary-d
34	86	95.6	20	AAW97234	Rat type ligand po
35	86	95.6	20	AAW95191	Bovine pituitary-d
36	86	95.6	20	AAW95175	Murine pituitary-d
37	86	95.6	20	AAW10350	Bovine oxytocin se
38	86	95.6	20	AAW10358	Rat oxytocin secre
39	86	95.6	20	AAV49301	1992 ligand peptid
40	86	95.6	20	AAV49302	1992 ligand peptid
41	86	95.6	20	AAW62519	Rat CRH releasin
42	86	95.6	20	AAW62527	Proactin releasin
43	86	95.6	20	AAW90994	Proactin releasin
44	86	95.6	20	AAW90996	Proactin releasin
45	86	95.6	20	AAW46594	Peptide PRR20 Itra

#### ALIGNMENTS

RESULT 1

AAW31394 standard; Peptide: 20 AA.

AC AAW31394;

XX 06-APR-1998 (first entry)

DT XX

XX Human type G protein-coupled receptor ligand fragment 4.

DE XX

XX G protein-coupled receptor; ligand binding; pharmaceutical;

KW modulator; pituitary; central nervous system; pancreas; prophylactic;

KW therapeutic agent.

XX

OS Homo sapiens.

XX

PN WO9724436-A2.

XX

PD 10-JUL-1997.

XX

XX 26-DEC-1996; 96WO-UP03821.

XX

XX 18-SEP-1996; 96UP-0246573.

PR 28-DEC-1995; 95UP-0343371.

PR 15-MAR-1996; 96UP-0059419.

PR 12-AUG-1996; 96UP-0211805.

XX

PA (TAKE ) TAKEDA CHEM IND LTD.

XX

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX WPI: 1997-363672/33.

DR N-PSDB; AAV02431.

XX Lligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAM97236 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. A central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC propylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharide. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA:  
  
Query Match 96.7%; Score 87; DB 18; Length 20;  
Best Local Similarity 88.2%; Pred. No. 9.9e-09;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 INPAMYXXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPWAYSGIRPVGRF 20  
  
RESULT 2  
AAM97236  
ID AAM97236 standard; peptide: 20 AA.  
XX  
AC AAM97236;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Human type ligand polypeptide fragment.  
XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacosgenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenioptathy; autoimmune disease;  
KW prolactinoma; infertility; amenorrhea; galactorrhoea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
KW Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.  
XX  
KW Homo sapiens.  
XX  
OS  
XX  
PN W09858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX

DR WPI; 1999-105614/09.  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 166; 241pp; English.  
XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonecyst cacosgenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenioptathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhoea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydralid mole, Irruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
SQ Sequence 20 AA:  
  
Query Match 96.7%; Score 87; DB 20; Length 20;  
Best Local Similarity 88.2%; Pred. No. 9.9e-09;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 INPAMYXXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPWAYSGIRPVGRF 20  
  
RESULT 3  
AAB10365  
ID AAB10365 standard; peptide: 20 AA.  
XX  
AC AAB10365;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS  
XX  
PN W0200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-452298/39.  
XX  
PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine





AAB90992  
ID AAB90992 standard; Peptide: 20 AA.  
XX  
AC AAB90992;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Prolactin releasing peptide SEQ ID NO:166.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KM blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PE 17-MAY-2000; 2000WO-US33576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI: 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
PS Disclosure: Page 244; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (iii) and a  
CC reactive group (ii) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (iv), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptide stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specifically as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 20 AA:  
XX  
Query Match 96.7%; Score 87; DB 22; Length 20;  
Best Local Similarity 88.2%; Pred. No. 9.9e-09;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 INPAMYXXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 7  
AAB31395  
ID AAB31395 standard; Peptide: 21 AA.  
XX  
AC AAB31395;  
XX

XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 5.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prolactin;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-034371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02432.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AAB31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trismus, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharide. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 21 AA:  
XX  
Query Match 96.7%; Score 87; DB 18; Length 21;  
Best Local Similarity 88.2%; Pred. No. 1e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 INPAMYXXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 8  
AAB10366  
ID AAB10366 standard; peptide: 21 AA.  
XX  
AC AAB10366;  
XX

DT	24-NOV-2000	(first entry)	
XX			
DE	Human oxytocin secretion promoting peptide SEQ ID NO: 36.		
XX			
KM	Human: oxytocin secretion promoter; G protein-coupled receptor protein;		
KM	treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;		
KM	caesarean section; artificial fertilization; galactostasis; goat; pig;		
KW	veterinary medicine; milk production.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200038704-A1.		
PD	06-JUL-2000.		
XX			
PE	22-DEC-1999; 99WO-JP07199.		
XX			
PR	25-DEC-1998; 98JP-0369585.		
XX			
PA	(TAKE ) TAKEDA CHEM IND LTD.		
PI	Matsumoto H, Kitada C, Hinuma S;		
XX			
DR	WPI: 2000-452298/39.		
XX			
PT	Physiologically-active polypeptide recognized as ligand by G		
PT	protein-coupled receptor protein, for promoting secretion of oxytocin,		
PT	as drugs for diseases relating to oxytocin secretion and in veterinary		
PT	medicine -		
XX			
PS	Disclosure; Page 63; 72pp; Japanese.		
XX			
CC	This invention describes a novel oxytocin secretion-regulating agent		
CC	which contains a ligand peptide or its salt for the G protein-coupled		
CC	receptor protein. It is useful in the form of drugs for ameliorating,		
CC	preventing and treating diseases relating to oxytocin secretion e.g.		
CC	weak pains and atonic bleeding, before and after expulsion of placenta,		
CC	uterine recovery failure, caesarean section, stoppage of artificial		
CC	fertilization or galactostasis and is also applicable in veterinary		
CC	medicine for promoting milk production in cow, goat and pig. This		
CC	sequence represents a human peptide which acts as an oxytocin secretion		
CC	promoter.		
CC			
SQ	Sequence 21 AA;		
	Query Match 96.7%; Score 87; DB 21; Length 21;		
	Best Local Similarity 88.2%; Pred. No. 1e-08;		
	Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;		
OY	1 INPAWYXXRGIRPVGRF 17		
DB	4 INPAWYASRGIRPVGRF 20		
RESULT 9			
AAAG62535			
ID	AAAG62535 standard; peptide; 21 AA.		
XX			
AC	AAAG62535;		
XX			
DT	24-AUG-2001 (first entry)		
XX			
DE	Human CRH releasing protein related peptide SEQ ID NO: 36.		
XX			
KM	Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;		
KM	analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;		
KW	Addison's disease; adrenal gland hyperfunction; obesity.		
XX			
OS	Homo sapiens.		
XX			
PN	WO200135984-A1.		

[illegible]

DR WPI: 1997-363672/33.  
DR N-PSDB; AAV02433.  
XX  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2: Page 186; 258pp; English.  
XX  
XX This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
XX Sequence 22 AA:  
S0

Query Match 96.7%; Score 87; DB 18; Length 22;  
Best Local Similarity 88.2%; Pred. No. 1.1e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPGVGRF 17  
| | | | | | | | | | | | | | | | | | | |  
DB 4 Inpawyasrgirpygrf 20

RESULT 11  
AAB10367  
ID AAB10367 standard; peptide: 22 AA.  
XX  
XX AAB10367:  
XX  
XX 24-NOV-2000 (first entry)  
XX  
XX Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
DE  
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
XX Homo sapiens.  
OS  
XX WO200038704-A1.  
XX  
XX 06-JUL-2000.  
XX  
XX 22-DEC-1999; 99WO-JP07199.  
XX  
XX 25-DEC-1998; 98JP-0369585.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
XX  
XX WPI: 2000-452298/39.  
XX  
XX Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT

PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
XX  
XX Disclosure; Page 64; 72pp; Japanese.  
XX  
XX  
XX This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
XX Sequence 22 AA:  
S0

Query Match 96.7%; Score 87; DB 21; Length 22;  
Best Local Similarity 88.2%; Pred. No. 1.1e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPGVGRF 17  
| | | | | | | | | | | | | | | | | | | |  
DB 4 Inpawyasrgirpygrf 20

RESULT 12  
AAG62536  
ID AAG62536 standard; peptide: 22 AA.  
XX  
XX AAG62536:  
XX  
XX 24-AUG-2001 (first entry)  
XX  
XX Human CRH releasing protein related peptide SEQ ID NO: 37.  
DE  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX Homo sapiens.  
OS  
XX WO200135984-A1.  
XX  
XX 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-JP08119.  
XX  
XX 18-NOV-1999; 99JP-0327900.  
XX  
XX 26-SEP-2000; 2000JP-0297073.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Kitada C, Matsumoto H, Hinuma S;  
XX  
XX WPI: 2001-35552/37.  
XX  
XX Use of G protein receptor ligand or peptide for controlling  
PT corticotropin releasing hormone secretion -  
PT  
XX  
XX Disclosure; Page 75; 90pp; Japanese.  
XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the

CC Invention.  
XX  
SQ Sequence 22 AA;

Query Match 96.7%; Score 87; DB 22; Length 22;  
Best Local Similarity 88.2%; Pred. No. 1.1e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRCIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 13  
AAW31391  
ID AAW31391 standard; Peptide; 31 AA.  
XX  
AC AAW31391;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 1.

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prolactin;  
KW therapeutic agent.

XX Homo sapiens.

XX MO9724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

XX 28-DEC-1995; 95JP-0343371.

XX 15-MAR-1996; 96JP-0059419.

XX 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;

XX WPI: 1997-363672/33.

XX N-PSDB: AAW02428.

XX Claim 2: Page 184; 258pp; English.

XX This sequence represents a peptide fragment from a novel human type

XX ligand polypeptide corresponding to amino acid residues 23 to 53 of the

XX sequence represented in AAW31390 and is used in an assay to monitor

XX ligand binding to the G protein-coupled receptor protein. Pharmaceutical

XX compositions containing this ligand may be used as a pituitary function

XX modulator, a central nervous system modulator or a pancreatic function

XX modulator. This ligand could have specific applications as a

XX prophyllactic or therapeutic agent for dementia, depression, hyperkinetic

XX syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,

XX trauma, growth hormone secretory disease, hyper- and polyphagia,

XX hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,

XX Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,

XX transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,

XX acute myocardial infarction, infertility, spinocerebellar degeneration,

XX bone fracture, trauma, atopic dermatitis, osteoporosis and/or

XX oligosacchara. Assays can also be developed to screen compounds which are

XX capable of altering the binding activity of the ligand affecting

XX activation of the G protein-coupled receptor protein.

XX  
SQ Sequence 31 AA;

Query Match 96.7%; Score 87; DB 18; Length 31;  
Best Local Similarity 88.2%; Pred. No. 1.6e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRCIRPVGRF 17  
||||| |||||||  
Db 15 INPAMYASRGIRPVGRF 31

RESULT 14  
AAW97235  
ID AAW97235 standard; peptide; 31 AA.  
XX  
AC AAW97235;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;

KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;

KW acromegaly; Chiari-Frommel syndrome; Argon-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
KW Irruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;

KW abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WO9858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fuji R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI: 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin

XX secretion or placental function, e.g. for treating menopausal

XX syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3: Page 159; 241pp; English.

XX The present sequence represents a human type ligand fragment. It

XX is used in the course of the invention. The specification describes

XX an agent for modulating prolactin secretion which comprises a

XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)

XX protein. The agents for promoting prolactin secretion can be used for

XX treating or preventing hypocoarismism, gonocyst cacosgenesis, menopausal

XX syndrome, euthyroid or hypometabolism. They can be used for promoting

XX lactation in a domestic mammal and as an aphrodisiac. The agents for

XX inhibiting prolactin secretion can be used for treating or preventing

XX pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,

XX prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,

XX acromegaly, Chiari-Frommel syndrome, Argon-del Castillo syndrome,

XX Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.

XX The inhibitory agents can also be used as contraceptives. The agents for

XX modulating placental function can be used for treating or preventing

XX choriocarcinoma, hydralid mole, Irruption mole, abortion, unfertilized fetus,

XX abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 31 AA;

Query Match 96.7%; Score 87; DB 20; Length 31;  
Best Local Similarity 88.2%; Pred. No. 1.6e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPVGRF 17  
||||| |||||||  
DB 15 INPAMYASRGIRPVGRF 31

RESULT 15  
AAM87615  
ID AAM87615 standard; Peptide: 31 AA.

XX AAM87615;

XX 29-MAR-1999 (first entry)

XX Human 19P2 ligand.

XX 19P2 ligand: G protein coupled receptor; pituitary;

KM prolactin releasing peptide; human; dementia; breast cancer;

KM therapy.

XX Homo sapiens.

XX EP887417-A2.

XX 30-DEC-1998.

XX 25-JUN-1998; 98EP-0111725.

XX 27-JUN-1997; 97JP-0172118.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Moriya T, Nishimura O, Suenaga M, Tanaka Y;

XX WPI: 1999-047884/05.

XX Producing a 19P2 pituitary G protein receptor ligand - by cleavage  
PT of a fusion protein, useful for preventing and treating dementia,  
PT breast cancer, renal failure and autoimmune disease

XX Claim 5; Page 35; 56pp; English.

XX This is the amino acid sequence of the human pituitary G  
CC protein-coupled receptor ligand 19P2L. A method suitable for  
CC commercial high-level production of 19P2L comprises expressing  
CC the ligand in host cells as a recombinant fusion protein e.g. with  
CC human basic fibroblast growth factor (see AAV83796-97) that has  
CC been modified to include an N-terminal cysteine residue. The  
CC ligand is released from the fusion by cyanylation followed by  
CC ammonolysis. 19P2L has prolactin secretion-stimulating and (at  
CC high doses) prolactin secretion-inhibiting properties. It can be  
CC used in the treatment and prevention of various diseases including:  
CC senile dementia, cerebrovascular dementia, and dementia associated  
CC with: neurological disorders (e.g. Alzheimer's disease, Parkinson's  
CC disease, Pick's disease, Huntington's disease), infectious diseases  
CC (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or  
CC toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism,  
CC intoxication by drugs, metal and organic compounds), tumorigenic  
CC diseases (e.g. brain tumour), traumatic diseases (e.g. chronic  
CC subarachnoidal haemorrhage, and other types of dementia, depression,  
CC hyperactive child syndrome (microcephalopathy) and disturbance of  
CC consciousness. It is also useful for prevention and treatment of  
CC diseases associated with prolactin hypo and hypersecretion  
CC respectively, including: hyperprolactinemia, pituitary adenoma,  
CC breast cancer, infertility, impotence and autoimmune disease  
CC (hypersecretion disorders), and seminal vesicle hypoplasia,

CC osteoporosis, menopausal syndrome and renal failure (hypersecretion  
CC disorders). The 19P2 polypeptide/amide is also useful as a test  
CC reagent for study of the prolactin secretory function or as a  
CC lactagogue in mammalian farm animals.

XX Sequence 31 AA;

Query Match 96.7%; Score 87; DB 20; Length 31;  
Best Local Similarity 88.2%; Pred. No. 1.6e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPVGRF 17  
||||| |||||||  
DB 15 INPAMYASRGIRPVGRF 31

Search completed: September 13, 2002, 09:18:35  
Job time: 500 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 Seconds  
(without alignments)  
3.218 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_4\_21  
Perfect score: 90  
Sequence: 1 INPAMYXXRGIRPYGRFX 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents, AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	96.7	20	3	US-09-105-678A-46
2	87	96.7	20	4	US-08-776-971-64
3	87	96.7	20	4	US-09-421-208-46
4	87	96.7	21	3	US-09-105-678A-47
5	87	96.7	21	4	US-08-776-971-65
6	87	96.7	21	4	US-09-421-208-47
7	87	96.7	22	3	US-09-105-678A-48
8	87	96.7	22	4	US-08-776-971-66
9	87	96.7	22	4	US-09-421-208-48
10	87	96.7	31	3	US-09-105-678A-9
11	87	96.7	31	3	US-09-105-678A-43
12	87	96.7	31	4	US-08-776-971-61
13	87	96.7	31	4	US-09-421-208-9
14	87	96.7	31	4	US-09-421-208-43
15	87	96.7	32	3	US-09-105-678A-44
16	87	96.7	32	4	US-08-776-971-62
17	87	96.7	32	4	US-09-421-208-44
18	87	96.7	33	3	US-09-105-678A-45
19	87	96.7	33	4	US-08-776-971-63
20	87	96.7	33	4	US-09-421-208-45
21	87	96.7	87	4	US-08-776-971-59
22	87	96.7	87	4	US-08-776-971-135
23	87	96.7	87	4	US-08-776-971-138
24	86	95.6	20	3	US-09-105-678A-34
25	86	95.6	20	3	US-09-105-678A-40
26	86	95.6	20	4	US-08-776-971-8
27	86	95.6	20	4	US-08-776-971-50

28	86	95.6	20	4	US-08-776-971-98	Sequence 98, Appl
29	86	95.6	20	4	US-09-421-208-34	Sequence 34, Appl
30	86	95.6	20	4	US-09-421-208-40	Sequence 40, Appl
31	86	95.6	21	3	US-09-105-678A-35	Sequence 35, Appl
32	86	95.6	21	3	US-09-105-678A-41	Sequence 41, Appl
33	86	95.6	21	4	US-08-776-971-9	Sequence 9, Appl
34	86	95.6	21	4	US-08-776-971-51	Sequence 51, Appl
35	86	95.6	21	4	US-09-421-208-35	Sequence 35, Appl
36	86	95.6	21	4	US-09-421-208-41	Sequence 41, Appl
37	86	95.6	22	3	US-09-105-678A-36	Sequence 36, Appl
38	86	95.6	22	3	US-09-105-678A-42	Sequence 42, Appl
39	86	95.6	22	4	US-08-776-971-10	Sequence 10, Appl
40	86	95.6	22	4	US-08-776-971-52	Sequence 52, Appl
41	86	95.6	22	4	US-09-421-208-42	Sequence 42, Appl
42	86	95.6	31	3	US-09-105-678A-7	Sequence 7, Appl
43	86	95.6	31	3	US-09-105-678A-8	Sequence 8, Appl
44	86	95.6	31	3	US-09-105-678A-31	Sequence 31, Appl
45	86	95.6	31	3	US-09-105-678A-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
; Sequence 46, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takao  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-46

Query Match 96.7%; Score 87; DB 3; Length 20;  
Best Local Similarity 88.2%; Pred. No. 1.2e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

```
QY 1 INPAWYXXRGIRPVGRF 17
    ||||| |||||
Db 4 INPAWYASRGIRPVGRF 20
```

```

RESULT 2
US-08-776-971-64
Sequence 64, Application US/08776971B
Patent No. 6228984
GENERAL INFORMATION:
APPLICANT: Hinuma, Shuji
            Habata, Yugo
            Kawamata, Yuji
            Hosoya, Masaki
            Fujii, Kyo
            Fukusumi, Shoji
            Kitada, Chieko
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
NUMBER OF SEQUENCES: 140
CORRESPONDENCE ADDRESS:
ADDRESSEE: DINE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
STREET: 130 Water Street
CITY: Boston
STATE: MA
COUNTRY: USA
ZIP: 02109
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/776, 971B
FILING DATE: 06-Feb-1997
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP96/03821
FILING DATE: 28-DEC-1996
APPLICATION NUMBER: JP 7/343371
FILING DATE: 28-DEC-1995
APPLICATION NUMBER: JP 8/59419
FILING DATE: 15-MAR-1996
APPLICATION NUMBER: JP 8/211805
FILING DATE: 12-AUG-1996
APPLICATION NUMBER: JP 8/246573
FILING DATE: 18-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: Conlin, David G.
REGISTRATION NUMBER: 27,026
REFERENCE/DOCKET NUMBER: 47176
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617-523-3400
TELEFAX: 617-523-6440
INFORMATION FOR SEQ ID NO: 64:
SEQUENCE CHARACTERISTICS:
LENGTH: 20 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
FRAGMENT TYPE: internal
SEQUENCE DESCRIPTION: SEQ ID NO: 64:
US-08-776-971-64

```

```

Query Match          96.7%  Score 87;  DB 4;  Length 20
Best Local Similarity 88.2%  Pred. NO. 1.2e-08;
Matches 15;  Conservative 0;  Mismatches 2;  Indels
Qy      1  INPAMYXXRGIRPVGRF 17
        |||||  |||||
Db      4  INPAMYASRGIRPVGRF 20

```

```

US-09-421-208-46
: Sequence 46, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESSES:
: ADDRESSEE: DIXE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 46:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 20 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
US-09-421-208-46

```

RESULT 4  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCE: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston



STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105, 678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 96.7%; Score 87; DB 3; Length 21;  
Best Local Similarity 88.2%; Pred. No. 1.3e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMYXXRCIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hiruma, Shuji  
Hiruma, Shuji  
Habata, Yugo  
Kawabata, Taji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 96.7%; Score 87; DB 4; Length 21;  
Best Local Similarity 88.2%; Pred. No. 1.3e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMYXXRCIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Suenaga, Masato  
Moriya, Takeo  
Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105, 678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 96.7%; Score 87; DB 4; Length 21;  
Best Local Similarity 88.2%; Pred. No. 1.3e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 96.7%; Score 87; DB 3; Length 22;  
Best Local Similarity 88.2%; Pred. No. 1.4e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 96.7%; Score 87; DB 4; Length 22;  
Best Local Similarity 88.2%; Pred. No. 1.4e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXRGIRPVGRF 17  
||||| |||||||  
Db 4 INPAMYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

```

; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-48

Query Match          96.7%; Score 87; DB 4; Length 22;
Best Local Similarity 88.2%; Pred. No. 1,4e-08;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMTXRGIRPVGRF 17
   ||||| |||||
DB 4 INPAMTXRGIRPVGRF 20

RESULT 10
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109

```

```

; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; MOLECULE TYPE: peptide
; US-09-105-678A-9

Query Match          96.7%; Score 87; DB 3; Length 31;
Best Local Similarity 88.2%; Pred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMTXRGIRPVGRF 17
   ||||| |||||
DB 15 INPAMTXRGIRPVGRF 31

RESULT 11
US-09-105-678A-43
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440

```

;; INFORMATION FOR SEQ ID NO: 43:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 31 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
US-09-105-678A-43

Query Match 96.7%; Score 87; DB 3; Length 31;  
Best Local Similarity 88.2%; Pred. No. 2e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMYXXRGIRPVGRF 17  
Db 15 INPAMYASRGIRPVGRF 31

RESULT 12  
US-08-776-971-61  
; Sequence 61, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
City: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 96.7%; Score 87; DB 4; Length 31;  
Best Local Similarity 88.2%; Pred. No. 2e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMYXXRGIRPVGRF 17  
Db 15 INPAMYASRGIRPVGRF 31

RESULT 13  
US-09-421-208-9  
; Sequence 9, Application US/09421208  
; Patent No. 6238561  
; GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
Moriya, Takeo  
Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
City: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 96.7%; Score 87; DB 4; Length 31;  
Best Local Similarity 88.2%; Pred. No. 2e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 INPAMYXXRGIRPVGRF 17  
Db 15 INPAMYASRGIRPVGRF 31

RESULT 14

```

US-09-421-208-43
: Sequence 43, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: PatentIn Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-09-421-208-43

Query Match 96.7%; Score 87; DB 4; Length 31;
Best Local Similarity 88.2%; Fred. No. 2e-08;
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0.

OY 1 INPANYXXRGIRPVGRF 17
| | | | | | | | | | | | | | | | | | | | | | | | | | |
DB 15 INPANYASRGIRPVGRF 31

RESULT 15
US-09-105-678A-44
: Sequence 44, Application US/09105678A
: Patent No. 6103882
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
:
:
:

```

```

      ZIP: 02109
      COMPUTER READABLE FORM:
      MEDIUM TYPE: Floppy disk
      COMPUTER: IBM PC compatible
      OPERATING SYSTEM: PC-DOS/MS-DOS
      SOFTWARE: PatentIn Release #1.0, Version #1.30
      CURRENT APPLICATION DATA:
      APPLICATION NUMBER: US/09/105,678A
      FILING DATE: 26-JUN-1998
      PRIOR APPLICATION DATA:
      APPLICATION NUMBER: JP 172118/1997
      FILING DATE: 27-JUN-1997
      ATTORNEY/AGENT INFORMATION:
      NAME: Conlin, David G.
      REGISTRATION NUMBER: 27,026
      REFERENCE/DOCKET NUMBER: 48466-342
      TELECOMMUNICATION INFORMATION:
      TELEPHONE: 617-523-3400
      TELEFAX: 617-523-6440
      INFORMATION FOR SEQ ID NO: 44:
      SEQUENCE CHARACTERISTICS:
      LENGTH: 32 amino acids
      TYPE: amino acid
      STRANDEDNESS:
      TOPOLOGY: linear
      MOLECULE TYPE: peptide
      US-09-105-678A-44

```

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:23:59 ; Search time 172.41 Seconds  
(without alignments)  
10.032 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_4\_21  
Perfect score: 90  
Sequence: 1 INPAWYXXRGIRPVGRFX 18

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	86	95.6	83	2	JC7607
2	46	51.1	333	2	H82852
3	46	51.1	790	2	T47959
4	43	47.8	220	2	C83292
5	43	47.8	719	2	S61046
6	42.5	47.2	443	2	T21499
7	42	46.7	433	2	H87660
8	42	46.7	226	2	A87664
9	41	45.6	338	2	T20100
10	41	45.6	342	2	B64395
11	41	45.6	347	2	H64371
12	41	45.6	476	2	G64720
13	41	45.6	476	2	G90629
14	41	45.6	476	2	G85480
15	41	45.6	476	2	AG0502
16	41	45.6	545	2	A87448
17	40	44.4	184	2	T35841
18	40	44.4	284	2	F71015
19	40	44.4	324	2	T35901
20	40	44.4	390	2	G82844
21	40	44.4	486	2	AF1174
22	40	44.4	486	2	AG1531
23	40	44.4	765	2	AE0418
24	39.5	43.9	779	2	AE0418
25	39	43.3	149	1	NEBP37
26	39	43.3	250	2	G83400
27	39	43.3	256	2	E95936
28	39	43.3	340	2	T20102
29	39	43.3	485	2	E83663

30	39	43.3	938	2	C84480	hypothetical prote
31	39	43.3	953	2	E82068	valyl-tRNA synthet
32	38.5	42.8	154	2	AH0264	conserved hypothet
33	38.5	42.8	374	2	G70947	hypothetical prote
34	38.5	42.8	831	2	S26675	DNA-directed DNA p
35	38.5	42.8	877	1	IJB0CN	N-cadherin precurs
36	38.5	42.8	906	1	IJH0CN	N-cadherin 2 precurs
37	38.5	42.8	906	1	IJH0CN	N-cadherin 2 precurs
38	38	42.2	127	2	B83157	hypothetical prote
39	38	42.2	158	2	F72725	probable ribosomal
40	38	42.2	167	2	AB2796	acetyltransferase
41	38	42.2	167	2	AB2796	hypothetical prote
42	38	42.2	178	2	D95026	ribosomal protein
43	38	42.2	178	2	D97897	50S ribosomal prot
44	38	42.2	194	2	D87357	hypothetical prote
45	38	42.2	238	2	H72646	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 95.6%; Score 86; DB 2; Length 83;  
Best local Similarity 88.2%; Pred. No. 8.2e-08;  
Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAWYXXRGIRPVGRF 17  
Db 36 INPAWYXXRGIRPVGRF 52

RESULT 2  
H82852  
hydroxybenzoate octaprenyltransferase XF0068 [imported] - Xylella fastidiosa (strain C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: H82852  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A55328 below  
A:Accession: H82852  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <SIM>  
A:Cross-references: GB:AE003860; GB:AE003849; NID:99104830; PIDN:AAF82881.1; GSPDB:GN A:Experimental source: strain 945C  
R:Stimpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R. Bioness, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Carraro, D.M.; Carre as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to Genbank, June 2000





A:Cross-references: EMBL:J270684; PIDD:CAA94603.1; GSPDB:GN00022; CESP:F28D1.8  
A:Experimental source: clone F28D1  
C:Genetics:  
A:Gene: CESP:F28D1.8  
A:Map position: 4  
A:Introns: 71/1; 103/3; 162/3; 215/2; 360/2

Query Match	47.28	Score 42.5	DB 2	Length 443
Best Local Similarity	37.58	Pred. NO. 15		
Matches 9, Conservative	3	Mismatches 3	Indels 9	Gaps 1

```
Qy      1 INPAWYXXR-----GIRPVG 15
          |||::|
Db      53 INPSWFDWRVSVSIDGHLGIWPIG 76
```

```

RESULT      7
H87660
peptidoglycan-binding protein, probable [imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence,revision 20-Apr-2001 #text,change 20-Apr-2001
C:Accession: H87660
R:Nierman, W.C.; Felblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.
B.; Iwab, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: A87249; MUID:21173698; PMID:11259647
A:Accession: H87660
A:Status: Preliminary
A:Molecule type: DNA
A:Residues: 1-433 <5MO>
A:Cross-references: GB:AE005673; NID:913425020; PIDN:AAK25284.1; GSPDB:GN00148
C:Genetics:
:Gene: CC3322

```

Query Match	46.7%	Score 42	DB 2	Length 433
Best Local Similarity	60.0%	Pred. NO. 18		
Matches 6	Conservative 2	Mismatches 2	Indels 0	Gaps 0

```

OY      3 PAWYXXRGIR 12
      |||: ||:|
Db      285 PAWWEARGVR 294

```

```

RESULT      8
AB87664
hypothetical protein CC3347 [Imported] - Caulobacter crescentus
C:Species: Caulobacter crescentus
C:Date: 20-Apr-2001 #sequence_revision 20-Apr-2001 #text_change 20-Apr-2001
C:Accession: AB87664
R:NIEMANN, W.C.; FELDBLYUM, T.V.; PAULSEN, I.T.; NELSON, K.E.; EISEN, J.; HELDEBERG, J.
B.; LAUB, M.T.; DEBOY, R.T.; DODGSON, R.J.; DURLIN, A.S.; GWINN, M.L.; HAFU, D.H.; KOLCH
N, J.; ERMOLAIEVA, M.; WHITE, O.; SALZBERG, S.L.; SHAPIRO, L.; VENTER, J.C.; FRASEE, C.M.
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001
A:Title: Complete Genome Sequence of Caulobacter crescentus.
A:Reference number: AB87249; MUID:21173698; PMID:11259647
A:Accession: AB87664
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-226 <5MO>
A:Cross-references: GB:AE005673; NID:g13425049; PIDN:AAK25309.1; GSPDB:GN00148
:Genetics:
:Gene: CC3347

```

Query Match	45.6%	Score 41;	DB 2;	Length 226;
Best Local Similarity	41.2%;	Pred. NO. 14;		
Matches	7;	Conservative	3;	Mismatches 7;
				Indels 0;
				Gaps 0.

```
Db      125 VNPDRSGRALRDVEQF 141
```

```

RESULT      9
T20100
hypothetical protein CS0C10.2 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C:Date: 15-Oct-1999 #revision 15-Oct-1999 #text_change 15-Oct-1999
C:Accession: T20100
C:McMurray, A.

```

A:Cross-references: EML:272505; PIDN:CAA96608.1; GSPDB:GMO0023; CESP:C50C10.2  
A:Experimental source: clone C50C10  
C:Genetics:  
A:Gene: CESP:C50C10.2  
A:Map position: 5  
A:Introns: 74/3, 144/3, 267/3

Query Match	45.68;	Score 41;	DB 2;	Length 338;
Best Local Similarity	53.88;	Pred. No. 21;		
Matches 7, Conservative	1;	Mismatches 5;	Indels 0;	Gaps 0

```

Oy      3 PAWYXXRGIRPVG 15
          | | | | | | |
Db      275 PYWYQILFIRPIG 281

```

RESULT 10  
B64395  
malic acid transport protein - Methanococcus jannaschii  
C:Species: Methanococcus jannaschii  
C:Date: 13-Sep-1996 #sequence\_revision 13-Sep-1996 #text\_change 21-Jul-2000  
C:Accession: B64395  
R:BULT, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blak  
; Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weinstock, R.G.; Merrick, J.M.; Glodex,  
rson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kaine, B.P.; Borodovsky, M.; Klek, H.P.; Fraser, C.M.; Smith, H.O.; Woese  
A:Title: Complete genome sequence of the methanogenic archaeon, Methanococcus jannasc  
A:Reference number: A64300; MUID:96337999  
A:Accession: B64395  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1342 <BULT>  
A:Cross-references: GB:U67521; GB:L77117; NID:g1591463; PIDN:AMB9875.1; PID:g1591472  
A:Genetics:  
A:Map position: REV685864-684836

Query Match	45.68	Score 41	DB 2	Length 342
Best Local Similarity	53.38	Pred. NO. 21		
Matches 8	Conservative 1	Mismatches 2	Indels 4	Gaps 1

```
QY      1 INPAWYXXRGIRPVG 15
          :||| | |||
Db      139 VNPGWY----IPVG 149
```

R: malle acid transport protein homolog - *Methanococcus jannaschl*  
H64371  
C/Species: *Methanococcus jannaschl*  
C/Date: 13-Sep-1996 #sequence.revision 13-Sep-1996 #text.change 21-Jul-2000  
C/Accession: H64371  
R:Bult, C.J.; White, O.; Olsen, G.J.; Zhou, L.; Fleischmann, R.D.; Sutton, G.G.; Blakes  
R: Reich, C.I.; Overbeek, R.; Kirkness, E.F.; Weissbrock, K.G.; Merrick, J.M.; Glöckl,  
R:

erson, J.D.; Sadow, P.W.; Hanna, M.C.; Cotton, M.D.; Roberts, K.M.; Hurst, M.A.  
Science 273, 1058-1073, 1996  
A:Authors: Kalne, B.P.; Borodovsky, M.; Klenk, H.P.; Fraser, C.M.; Smith, H.O.; Woese, C.  
A:Title: Complete genome sequence of the methanogenic archaeon, *Methanococcus jannaschii*.  
A:Reference number: A64300; MUID:96337999  
A:Accession: H64371  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-347 <BUIL>  
A:Cross-references: GB:067506; GB:L77117; NID:g1591274; PIDN:AAB8567.1; PID:g1591283; T  
C:Genetics:  
I:Map position: FOR511924-512967

Query Match	45.6%	Score 41:	DB 2:	Length 347:
Best Local Similarity	53.3%	Pred. No. 21:		
Matches	8:	Conservative	1:	Mismatches 2:
				Indels 4:
				Gaps 1:
Qy	1	INPARYXXRGIRPVG	15	
	:			
Db	146	VNPGMY----	IPVG 156	

RESULT 12  
664720  
Probable amino acid transport protein yaaJ, sodium-dependent - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
C:Accession: 664720  
R:Baltner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; CC  
A.: Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: 664720  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <BLAT>  
A:Cross-references: GB:AE000111; GB:U00096; NID:g1786181; PIDN:AAC73118.1; PID:g1786188;  
A:Experimental source: strain K-12, substrain MC1555  
C:Genetics:  
A:Gene: yaaJ  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein  
C:Keywords: amino acid transport; transmembrane protein  
F:10-26/Domain: transmembrane #status predicted <TM1>  
F:91-107/Domain: transmembrane #status predicted <TM2>  
F:142-158/Domain: transmembrane #status predicted <TM3>  
F:178-194/Domain: transmembrane #status predicted <TM4>  
F:208-224/Domain: transmembrane #status predicted <TM5>  
F:303-319/Domain: transmembrane #status predicted <TM6>  
F:349-365/Domain: transmembrane #status predicted <TM7>  
F:391-407/Domain: transmembrane #status predicted <TM8>  
F:414-430/Domain: transmembrane #status predicted <TM9>

```

Query Match          45.6%; Score 41; DB 2; Length 476;
Best Local Similarity 66.7%; Pred. No. 29;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY      3 PAMVXXRCI 11
        |||||
        |||
Db       129 PAMVWRCI 137

RESULT 13
G90629
Probable inner membrane transport protein EGS0007 [Imported] - Escherichia coli (strain
C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 03-Aug-2001
C:Accession: G90629
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.
Nucleic
Acids Res. 8, 11-22, 2001

```

A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90629  
A:Status: Preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA03430.1; PID:G1335963; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: ECs007  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

	Query Match.	45.6%;	Score 41;	DB 2;	Length 476;
	Best Local Similarity	66.7%;	Pred. No. 29;		
Matches	6;	Conservative	1;	Mismatches	2; Indels 0; Gaps
OY	3 PAMXXXXGI 11				
	11:				
Db	129 PAMIMARGL 137				

RESULT 14  
GB5480  
Inner membrane transport protein [imported] - *Escherichia coli* (strain O157:H7, substid  
C.Species: *Escherichia coli*  
C.Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-sep-2001  
C.Accession: GB54801  
R.Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayn  
Iller, L.; Grobbeck, E.J.; Davis, N.W.; Lim, A.; Diallanita, E.; Potamoustis, K.; Apoda  
Nature 409, 529-533, 2001  
A.Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A.Reference number: AB5480; MUID:21074935; PMID:11206551  
A.Accession: GB5480  
A.Status: Preliminary  
A.Molecule type: DNA  
A.Residues: 1-476 <SMO>  
A.Cross-references: GB:AB005174; NTD:g12512682; PIDN:AA654307.1; GSPDB:GN00145; UMGP:  
A.Experimental source: strain O157:H7, substrain EDL933  
C.Genetics:  
A.Gene: yaaU  
C.Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match	45.6%	Score 41:	DB 2:	Length 476;
Best Local Similarity	66.7%	Pred. No. 29;		
Matches	6:	Conservative	1:	Mismatches 2: Indels 0: Gaps 0:
QY	3	PAWYXXRCI	11	
Db	129	PAWYMARGL	137	

RESULT 15  
 AG0502  
 Probable amino-acid transport protein SRY0006 [imported] - *Salmonella enterica* subsp.  
 C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh1  
 A:Note: this species has also been called *Salmonella typhi*  
 C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AG0502  
 P:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
 th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar  
 , S.; Moulé, S.; O'Gaora, P.  
 Nature 413, 848-852, 2001  
 A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
 A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* se  
 A:Reference number: AB0502; PMID:11677608  
 A:Accession: AG0502  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-476 <FAR>  
 A:Cross-references: GB:AL513382; PIDN:CAD01159.1; PID:gl6501289; GSPDB:GN00176  
 ;Genetics:

A:Gene: STY0006  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 45.6%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 29;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 3 PAWYXXRGI 11  
111111  
Db 129 PAWYMARGL 137

Search completed: September 13, 2002, 09:23:59  
Job time: 774 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:45 ; Search time 80.21 Seconds

(without alignments)  
8.689 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_4\_21

Perfect score: 90

Sequence: 1 INPAMWXXRGIRPVGRFX 18

Scoring table: BLOSUM62

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	87	96.7	87	1	PRRP_HUMAN
2	86	95.6	83	1	PRRP_RAT
3	86	95.6	98	1	PRRP_BOVIN
4	44	48.9	676	1	EXL1_HUMAN
5	43	47.8	719	1	NRPI_YEAST
6	42	46.7	402	1	EX7L_SPROCO
7	41	45.6	342	1	Y576_METUA
8	41	45.6	347	1	Y576_METUA
9	41	45.6	476	1	YAAO_ECOLI
10	39	43.3	149	1	ENRN_BPT7
11	39	43.3	485	1	SYV_BACHD
12	39	42.8	953	1	SYV_VIRCH
13	38.5	42.8	831	1	DPOI_THEFL
14	38.5	42.8	877	1	CAD2_BOVIN
15	38.5	42.8	906	1	CAD2_HUMAN
16	38.5	42.8	906	1	CAD2_MOUSE
17	38.5	42.8	158	1	RL15_AERPE
18	38	42.2	906	1	MTNA_DROME
19	38	42.2	637	1	MTNA_DROME
20	38	42.2	951	1	SYV_ECOLI
21	38	42.2	1723	1	AIM1_HUMAN
22	38	42.2	2832	1	NDVB_RHIME
23	37	41.1	179	1	IN37_SPIOL
24	37	41.1	344	1	OP51_LIMPO
25	37	41.1	376	1	OP52_LIMPO
26	37	41.1	376	1	OP52_LIMPO
27	37	41.1	385	1	Y4EB_ECOLI
28	37	41.1	640	1	Y4CD_RHIN
29	37	41.1	788	1	RECA_HAEIN
30	36	40.0	241	1	Y63D_YEAST
31	36	40.0	329	1	Y493_MYCTU
32	36	40.0	334	1	YNHG_ECOLI
33	36	40.0	459	1	VL2_HPVA

34	36	40.0	459	1	VL2_HPVA	P03106 human papill
35	36	40.0	994	1	SVV_XYLEA	O9P112 xylella fas
36	36	40.0	2163	1	BRR2_YEAST	P32639 saccharomyc
37	35.5	39.4	349	1	YHHT_ECOLI	P37622 escherichia
38	35	38.9	135	1	RYK16_SPIOG	P06510 splirodela o
39	35	38.9	289	1	LYTB_AQUAE	O67625 aquifex aco
40	35	38.9	379	1	Y528_STNY3	O55318 synechocyst
41	35	38.9	383	1	CYCR_CHRVI	O82947 chromatiu
42	35	38.9	390	1	RM03_YEAST	P36516 saccharomyc
43	35	38.9	391	1	THIL_ZOORA	P07097 zoogloea ra
44	35	38.9	422	1	GP72_CANFA	O9tc9 canis fam11
45	35	38.9	423	1	GP72_HUMAN	O9nym4 homo sapien

## ALIGNMENTS

RESULT	1	STANDARD	PRT	87 AA
PRRP_HUMAN				
ID	P81277			
AC	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP31, Proactin-releasing peptide PRRP20].			
DE	PRH.			
GN	Homo sapiens (Human).			
OS	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.			
OX	NCBI_TaxID=9606;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=98268781; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Hinuma S., Kikada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RA	"A proactin-releasing peptide in the brain.";			
RT	Nature 393:272-276(1998).			
RT	TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.			
RL	Regul. Pept. 83:1-10(1999).			
CC	-1- FUNCTION: Stimulates proactin (PRU) release and regulates the expression of proactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.			
CC	-1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.			
CC	*****			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	*****			
DR	EMBL, AB015419; BAA29027.1; -			
DR	MM: 602663; -			
KW	Hormone; Amidation; Signal.			
FT	SIGNAL			
FT	PEPTIDE			
FT	PEPTIDE			
FT	MOD RES			
FT	SEQUENCE			
FT	BY SIMILARITY.			
FT	PROACTIN-RELEASING PEPTIDE PRRP31.			
FT	PROACTIN-RELEASING PEPTIDE PRRP20.			
FT	AMIDATION (G-54 PROVIDE AMIDE GROUP).			
FT	229A2P35FC9C981B CXC64;			

Query Match 96.7%; Score 87; DB 1; Length 87;  
 Best Local Similarity 88.2%; Pred. NO. 1e-08;  
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPVGRF 17  
 ||||| |||||  
 DB 37 INPAMYASRGIRPVGRF 53

## RESULT 2

PRRP\_RAT STANDARD; PRT: 83 AA.

AC P81276;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: proactin-releasing peptide PRRP1; proactin-releasing peptide PRRP2].  
 GN PRH.

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;

RP SEQUENCE FROM N.A.

RC TISSUE-Brain;  
 RX MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;

RT "A proactin-releasing peptide in the brain."  
 RL Nature 393:272-276(1998).

RN TISSUE SPECIFICITY.

RP PubMed=10498338.

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.;

RT "Tissue distribution of proactin-releasing peptide (PRP) and its receptor."  
 RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR1 EMBL; AB015418; BAA29026.1; Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT SIGNAL 1 21 BY SIMILARITY.

FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP1.

FT MOD\_RES 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.

FT MOD\_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).

SO SEQUENCE 83 AA; 9215 MW; DDC75A264EBE4F29 CRC64;

Query Match 95.6%; Score 86; DB 1; Length 83;  
 Best Local Similarity 88.2%; Pred. NO. 1.4e-08;  
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPVGRF 17  
 ||||| |||||  
 DB 36 INPAMYTGIRPVGRF 52

RESULT 3  
 PRRP\_BOVIN STANDARD; PRT: 98 AA.

AC P81264;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: proactin-releasing peptide PRRP1; proactin-releasing peptide PRRP2].  
 GN PRH.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;

RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RC TISSUE-Brain;  
 RX MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;

RT "A proactin-releasing peptide in the brain."  
 RL Nature 393:272-276(1998).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; AB015417; BAA29025.1; Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT SIGNAL 1 22

FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP1.

FT MOD\_RES 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.

FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).

SO SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 95.6%; Score 86; DB 1; Length 98;  
 Best Local Similarity 88.2%; Pred. NO. 1.7e-08;  
 Matches 15; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIRPVGRF 17  
 ||||| |||||  
 DB 37 INPAMYAGIRPVGRF 53

## RESULT 4

EX1L\_HUMAN STANDARD; PRT: 676 AA.

AC Q92935;  
 DT 15-JUL-1998 (Rel. 36, Created)

DT 15-JUL-1998 (Rel. 36, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Exostosin-like 1 (Exostosin-L) (Multiple exostosin-like protein).

GN EX1L OR EXT1.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97189339; PubMed=9037597;

RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
 RT "Identification and localization of the gene for EXTL1, a third member  
 RT of the multiple exostosose gene family.";  
 RL Genome Res. 7:10-16(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;  
 RT "Mutations of the EXT genes in hereditary multiple exostososes in  
 RT Chinese.";  
 RL Submitted (Aug-1998) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Wuyts W., Spicker N., Van Roy N., De Paeppe A., De Bouille K.,  
 RA Willems P.J., Van Hul W., Versteeg R., Speleman F.;  
 RT "Refined physical mapping and genomic structure of the EXTL1 gene.";  
 RL Submitted (May-1999) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE EXSTOSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: U67191; AAC51141.1; -  
 DR EMBL: AF083633; AAD02840.1; -  
 DR EMBL: AF083623; AAD02840.1; JOINED.  
 DR EMBL: AF083624; AAD02840.1; JOINED.  
 DR EMBL: AF083625; AAD02840.1; JOINED.  
 DR EMBL: AF083626; AAD02840.1; JOINED.  
 DR EMBL: AF083627; AAD02840.1; JOINED.  
 DR EMBL: AF083628; AAD02840.1; JOINED.  
 DR EMBL: AF083629; AAD02840.1; JOINED.  
 DR EMBL: AF083630; AAD02840.1; JOINED.  
 DR EMBL: AF083631; AAD02840.1; JOINED.  
 DR EMBL: AF083632; AAD02840.1; JOINED.  
 DR EMBL: AF153980; AAF73172.1; -  
 DR EMBL: AF153991; AAF73172.1; JOINED.  
 DR MIM: 601738; -  
 DR InterPro: IPR004263; Exostosin.  
 DR Pfam: PF03016; Exostosin: 1.  
 DR Antl-oncogene: Multigene family: Transmembrane: Signal-anchor  
 DR TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 DR CARBOHYD 269 269 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 DR SEQUENCE 676 AA; 74673 MW; B5E06A8762E5633 CRC64;  
 SQ  
 Query Match 48.9%; Score 44; DB 1; Length 676;  
 Best Local Similarity 53.3%; Pred. No. 3.4;  
 Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 Oy 3 PAMYXXRGIRPVGRF 17  
 Db 405 PEYILOGGSRPEGRF 419

OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-AH22;  
 RX MEDLINE-93247548; PubMed-8483449;  
 RA Wehner E.P., Rao E., Brendel M.;  
 RT "Molecular structure and genetic regulation of SFA, a gene  
 RT responsible for resistance to formaldehyde in Saccharomyces  
 RL cerevisiae, and characterization of its protein product.";  
 RL Mol. Gen. Genet. 237:351-358(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-S288C;  
 RA Pohl T.M.;  
 RL Submitted (Nov-1995) to the EMBL/Genbank/DBJ databases.  
 CC -1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.  
 CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X68020; CAA48159.1; -  
 DR EMBL: Z67750; CAA91579.1; -  
 DR EMBL: Z74215; CAA98741.1; -  
 DR PIR: S31139; S31139.  
 DR HSSP: P04170; GRKN.  
 DR SGD: S0002326; NRPL.  
 DR InterPro: IPR000504; RRM.  
 DR InterPro: IPR001876; Znf-RanBP.  
 DR Pfam: PF00076; RRM\_1.  
 DR Pfam: PF00641; Zf-RanBP; 2.  
 DR SMART: SM00360; RRM; 1.  
 DR SMART: SM00547; Znf-RBZ; 2.  
 DR PROSITE: PS00102; RRM; 1.  
 DR PROSITE: PS00030; RRM\_RNP\_1; FALSE\_NEG.  
 DR PROSITE: PS01358; ZF\_RANBP2\_1; 2.  
 DR PROSITE: PS50199; ZF\_RANBP2\_2; 2.  
 DR Nuclear protein; zinc-finger; RNA-binding; Repeat.  
 DR DOMAIN 226 332  
 DR ZN\_FING 355 384 RNA-BINDING (RRM).  
 DR ZN\_FING 581 610 RANBP2-TYPE 1.  
 DR ZN\_FING 490 564 RANBP2-TYPE 2.  
 DR DOMAIN 493 493 ASN-RICH.  
 DR CONFLICT 493 493 I -> N (IN REF. 1).  
 DR SEQUENCE 719 AA; 79239 MW; ADA9BC09FD582669 CRC64;  
 SQ  
 Query Match 47.8%; Score 43; DB 1; Length 719;  
 Best Local Similarity 50.0%; Pred. No. 5.4;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;  
 Oy 4 AMYXXRGIRPVG 15  
 Db 244 SWFTQYGVAPVG 255

RESULT 5  
 NRPL\_YEAST  
 ID NRPL\_YEAST STANDARD: PRT: 719 AA.  
 AC P32770: 012228;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Asparagine-rich protein (ARP protein).  
 DE NRPL OR ARP1 OR ARP OR YDL167C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;

RESULT 6  
 EXTL\_STRCO  
 ID EXTL\_STRCO STANDARD: PRT: 402 AA.  
 AC Q9FBM3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
 DE (Exonuclease VII large subunit).  
 GN XSEA OR SCK7.29C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA STRAIN-A3(2):  
RA Seeger K.J., Harris D., Cerdano A.M., Parkhill J., Barrell B.G.,  
RA Rajandream M.A.,  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Exonuclease activity cleavage in either 5' to 3'  
CC or 3' to 5'-direction to yield 5'-phosphomononucleotides.  
CC -1- SUBUNIT: HETERODIMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: AL391754; CAC05901.1;  
CC InterPro: IPR003753; Exonuc\_VII\_L.  
CC InterPro: IPR003753; Exonuc\_VII\_L.  
CC Pfam: PF03601; Exonuc\_VII\_L; 1.  
CC Pfam: PF03601; Exonuc\_VII\_L; 1.  
CC Hydrolyase; Nuclease; Exonuclease.  
CC SEQUENCE 402 AA; 43682 MW; 145929A8372B4E08 CRC64;  
SO  
Query Match 46.7%; Score 42; DB 1; Length 402;  
Best Local Similarity 42.9%; Pred. No. 4.6;  
Matches 9; Conservative 1; Mismatches 3; Indels 8; Gaps 1;  
OY 3 PAMVXXRG-----IRPVG 15  
ID Y576\_METUA STANDARD; PRT; 342 AA.  
AC O58172;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein M0762.  
GN M0762.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merriam J.M., Glöck A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RA jannaschii.";  
RT Science 273:1058-1073(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC

CC -1- SIMILARITY: STRONG, TO M. JANNASCHII M0576 AND TO S. POMBE MALATE  
CC PERMEASE (MAE1).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
-----  
CC EMBL: U67521; AAB98753.1;  
CC TIGR; M0762;  
DR TIGR; M0762;  
KW Hypothetical protein; Transmembrane; Transport; Complete proteome.  
FT TRANSMEM 8 28 POTENTIAL.  
FT TRANSMEM 39 59 POTENTIAL.  
FT TRANSMEM 79 99 POTENTIAL.  
FT TRANSMEM 108 128 POTENTIAL.  
FT TRANSMEM 142 162 POTENTIAL.  
FT TRANSMEM 175 195 POTENTIAL.  
FT TRANSMEM 207 227 POTENTIAL.  
FT TRANSMEM 242 262 POTENTIAL.  
FT TRANSMEM 276 296 POTENTIAL.  
FT TRANSMEM 304 324 POTENTIAL.  
SO SEQUENCE 342 AA; 39534 MW; 08FEC3E2C4955D8 CRC64;  
OY 1 INPAMVXXRGIRPVG 15  
ID Y576\_METUA STANDARD; PRT; 347 AA.  
AC O57996;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein M0576.  
GN M0576.  
OS Methanococcus jannaschii.  
OC Archaea; Euryarchaeota; Methanococcales; Methanococcaceae;  
OC Methanococcus.  
OX NCBI\_TaxID=2190;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-JAL-1 / DSM 2661 / ATCC 43067;  
RX MEDLINE=96337999; PubMed=8688087;  
RA Bult C.J., White O., Olsen G.J., Zhou L., Fleischmann R.D.,  
RA Sutton G.G., Blake J.A., Fitzgerald L.M., Clayton R.A., Gocayne J.D.,  
RA Kerlavage A.R., Dougherty B.A., Tomb J.-F., Adams M.D., Reich C.I.,  
RA Overbeek R., Kirkness E.F., Weinstock K.G., Merriam J.M., Glöck A.,  
RA Scott J.L., Geoghegan N.S.M., Weidman J.F., Fuhmann J.L., Nguyen D.,  
RA Uterback T.R., Kelley J.M., Peterson J.D., Sadow P.W., Hanna M.C.,  
RA Cotton M.D., Roberts K.M., Hurst M.A., Kaine B.P., Borodovsky M.,  
RA Klenk H.-P., Fraser C.M., Smith H.O., Woese C.R., Venter J.C.;  
RA "Complete genome sequence of the methanogenic archaeon, Methanococcus  
RA jannaschii.";  
RT Science 273:1058-1073(1996).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
CC -1- SIMILARITY: STRONG, TO M. JANNASCHII M0762 AND TO S. POMBE MALATE  
CC PERMEASE (MAE1).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC





CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

DR EMBL; V01127; CAA24345.1; -  
DR EMBL; V01146; CAA24402.1; -  
DR PIR; A00785; NEBP37.  
DR PIR; S42301; S42301.  
KW Hydrolyase; Nuclease; Endonuclease.  
SQ SEQUENCE 149 AA; 17172 MW; D092AA28E3743BC1 CRC64;

Query Match 43.3%; Score 39; DB 1; Length 149;  
Best Local Similarity 57.1%; Pred. No. 5.9;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 4 AMYXXRGIRPVGR 17  
| : ||| |||  
DB 2 AGYGAKGIRKVGAF 15

## RESULT 11

SVL\_BACHD STANDARD; PRT; 485 AA.  
AC 09KGF6: 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE GlutamyL-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)  
DE (GluRS).  
GN GLTX OR BH0109.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RA MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +  
diphosphate + L-glutamyl-tRNA(Glu).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

CC EMBL; AP001507; BAB03828.1; -  
DR InterPro: IPR000924; tRNA-synt\_1c.  
DR InterPro: IPR001412; tRNA-synt\_1i.  
DR Pfam; PF00749; tRNA-synt\_1c; 1.  
DR PRINTS; PR00987; TRNASYNTGLOU.  
KW AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.

KW Complete proteome.  
FT SITE 11 "HIGH" REGION.  
FT SITE 252 "KMSK" REGION.  
FT BINDING 255 ATP (BY SIMILARITY).

FT SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;  
SQ

Query Match 43.3%; Score 39; DB 1; Length 485;

Best Local Similarity 58.3%; Pred. No. 19;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 6 YXXRGIRPVGR 17  
| : ||| |||  
DB 146 YEAKGIRPVGR 157

## RESULT 12

SVL\_VIBCH STANDARD; PRT; 953 AA.  
AC 09KPF3: 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Valyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (ValRS).  
GN VALS OR VC2503.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_TaxID=666;  
[1]

RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RA MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Uitterlinden R.D., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;

RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae.";  
RL Nature 406:477-483(2000).

CC -1- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate  
+ L-valyl-tRNA(Val).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@lsb-sib.ch](mailto:license@lsb-sib.ch)).

CC EMBL; AE004320; AAP95645.1; -  
DR HSSP; P96142; IGAX.  
DR TIGR; VC2503; -

DR InterPro: IPR002300; tRNA-synt\_1a.  
DR InterPro: IPR001412; tRNA-synt\_1i.  
DR InterPro: IPR002303; tRNA-synt\_val.  
DR Pfam; PF00133; tRNA-synt\_1i; 1.  
DR PRINTS; PR00986; TRNASYNTVAL.

KW AMINOACYL-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
DR PROSITE; PS00178; AA\_TRNA\_LIGASE\_I; 1.  
KW Complete proteome.  
FT SITE 42 "HIGH" REGION.  
FT SITE 554 "KMSK" REGION.  
FT BINDING 557 ATP (BY SIMILARITY).

FT SEQUENCE 953 AA; 108170 MW; D93471A33CP4F69C CRC64;  
SQ

Query Match 43.3%; Score 39; DB 1; Length 953;  
Best Local Similarity 57.1%; Pred. No. 37;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 3 PAMYXXRGIRPVGR 16  
||| : | |||  
DB 436 PAMYDEGQNVFGR 449

RESULT 13  
ID DPOL\_THEFL STANDARD: PRT: 831 AA.  
AC P30313:  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE DNA polymerase I, thermostable (EC 2.7.7.7) (TFL polymerase 1).  
GN POLA OR POL.  
OS Thermus aquaticus (subsp. flavus).  
OC Bacteria; Thermus/deinococcus group; Thermus group; Thermus.  
OX NCBI\_TaxID=274;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ACM-B-1257;  
RX MEDLINE=93087201; PubMed=1454544;  
RA Akmetzhanov A.A., Vakhitov V.A.;  
RT Molecular cloning and nucleotide sequence of the DNA polymerase gene  
from Thermus flavus.  
RL Nucleic Acids Res. 20:5839-5839(1992).  
CC -1- CATALYTIC ACTIVITY: N deoxynucleoside triphosphate - N diphosphate  
+ (DNA)(N).  
CC -1- SIMILARITY: BELONGS TO DNA POLYMERASE TYPE-A FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: X6105; CAA46900.1; -  
DR PIR: S24929; S24929.  
DR PIR: S26675; S26675.  
DR HSSP: P19821; ITAO.  
DR InterPro: IPR002421; 5\_3\_exonuclease.  
DR InterPro: IPR002298; DNA\_pol.  
DR InterPro: IPR001098; DNA\_pol\_A.  
DR InterPro: IPR000513; Exo\_N.I.  
DR InterPro: IPR003583; HHH\_1.  
DR InterPro: IPR003584; HHH\_2.  
DR InterPro: IPR001532; XPG\_1.  
DR Pfam: PF01367; 5\_3\_exonuclease\_1.  
DR Pfam: PF02739; 5\_3\_exonuclease\_N\_1.  
DR Pfam: PF00476; DNA\_pol\_A\_1.  
DR PRINTS: PR00868; DNAPOL1.  
DR SMART: SM00475; S3EXOC.1.  
DR SMART: SM00278; HHH1.1.  
DR SMART: SM00279; HHH2.1.  
DR SMART: SM00482; POLAC.1.  
DR SMART: SM00485; XRGH.1.  
DR PROSITE: PS00447; DNA\_POLYMERASE\_A\_1.  
KW Transference; DNA-directed DNA polymerase; DNA replication; DNA repair;  
KW DNA-binding.  
FT DOMAIN 409 831 POLYMERASE (BY SIMILARITY).  
FT SEQUENCE 831 AA; 93783 MW; 96F933CEFA3CA536D CRC64;  
Query Match 42.8%; Score 38.5; DB 1; Length 831;  
Best Local Similarity 57.1%; Pred. No. 39;  
Matches 8; Conservative 1; Mismatches 4; Indels 1; Gaps 1;  
QY 1 INPAM-YXXRGIRP 13  
Db 162 ITPAMLYEYGLRP 175  
RESULT 14  
ID CAD2\_BOVIN STANDARD: PRT: 877 AA.  
AC P19534;  
Query Match 42.8%; Score 38.5; DB 1; Length 877;  
Best Local Similarity 50.0%; Pred. No. 42;  
Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
QY 3 PAN---YXXRGIRPGRF 17

DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-FEB-1991 (Rel. 17, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neural-cadherin precursor (N-cadherin) (cadherin-2) (Fragment).  
GN CDH2.  
OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.  
OX NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90360979; PubMed=2390969;  
RA Liaw C.W., Cannon C., Power M.D., Kiboneke P.K., Rubin L.L.;  
RT Identification and cloning of two species of cadherins in bovine  
endothelial cells.  
RL EMBO J. 9:2701-2708(1990).  
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
NEURONAL RECOGNITION MECHANISM.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC EMBL: X53615; CAA37677.1; -  
DR PIR: S11693; IIBOCN.  
DR HSSP: P15116; INCI.  
DR InterPro: IPR002126; Cadherin.  
DR InterPro: IPR000233; Cadherin\_C-term.  
DR Pfam: PF00028; cadherin; 5.  
DR Pfam: PF01049; Cadherin\_C-term; 1.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA\_5.  
DR PROSITE: PS00232; CADHERIN\_1; 3.  
DR PROSITE: PS00268; CADHERIN\_2; 5.  
KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.  
FT CELL\_TER 1 1  
FT PROPEP <1 130  
FT CHAIN 131 877  
FT DOMAIN 131 695  
FT TRANSMEM 696 717  
FT DOMAIN 718 877  
FT DOMAIN 131 238  
FT DOMAIN 239 363  
FT DOMAIN 364 468  
FT DOMAIN 469 574  
FT DOMAIN 575 685  
FT DOMAIN 685 849  
FT CARBOHYD 161 161  
FT CARBOHYD 244 244  
FT CARBOHYD 296 296  
FT CARBOHYD 373 373  
FT CARBOHYD 543 543  
FT CARBOHYD 622 622  
FT CARBOHYD 663 663  
FT SEQUENCE 877 AA; 96845 MW; 441B829DE871A249 CRC64;  
Query Match 42.8%; Score 38.5; DB 1; Length 877;  
Best Local Similarity 50.0%; Pred. No. 42;  
Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
QY 3 PAN---YXXRGIRPGRF 17

DB 385 PAMNAVRIISGDPAGR 402

RESULT 15

CAD2\_HUMAN

ID CAD2\_HUMAN STANDARD; PRT: 906 AA.

AC P19022; Q14923;

DT 01-NOV-1990 (Rel. 16, Created)

DT 01-OCT-1996 (Rel. 34, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).

GN CDH2 OR CDH9 OR NCAD.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

OX NCBI\_TaxID=9606;

LN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE-91016946; PubMed-2216790;

RA Reid R.A., Hemperly J.J.;

RT "Human N-cadherin: nucleotide and deduced amino acid sequence.";

RL Nucleic Acids Res. 18:5896-5896(1990).

RN [2]

RP REVISIONS TO 341: 699 AND 705.

RA Reid R.A.;

RL Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.

RN [3]

RP SEQUENCE FROM N.A.

RX MEDLINE-92363956; PubMed-1500442;

RA Salomon D., Ayala O., Patel-King R., Hynes R.O., Gelber B.;

RT "Extrajunctional distribution of N-cadherin in cultured human endothelial cells.";

RL J. Cell Sci. 102:7-17(1992).

RN [4]

RP SEQUENCE OF 160-906 FROM N.A.

RX MEDLINE-90347462; PubMed-2384753;

RA Walsh F.S., Barton C.H., Putt W., Moore S.E., Kelsell D.;

RT "N-cadherin gene maps to human chromosome 18 and is not linked to the E-cadherin gene.";

RL J. Neurochem. 55:805-812(1990).

RN [5]

RP SEQUENCE OF 1-20 FROM N.A.

RX MEDLINE-95048366; PubMed-7959764;

RA Wallis J.A., Fox M., Walsh F.S.;

RT "Structure of the human N-cadherin gene: YAC analysis and fine genomic mapping to 18q11.2.";

RL Genomics 22:172-179(1994).

CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS. THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN NEURONAL RECOGNITION MECHANISM.

CC -1- SUBCELLULAR LOCATION: Type I membrane protein.

CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.

CC

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC

CC EMBL: X57548; CAA40773.1; -

DR EMBL: X54315; CAA38213.1; -

DR EMBL: S42303; AAB22854.1; -

DR EMBL: M34064; AAA03236.1; -

DR EMBL: Z27420; CAA81799.1; -

DR PIR: A38870; IJHUCN.

DR HSSP: P15116; INCIJ.

DR MIM: 114020; -

DR InterPro: IPR002126; Cadherin.

DR InterPro: IPR00233; Cadherin\_C\_term.

DR Pfam: PF00028; cadherin\_5.

DR Pfam: PF01049; cadherin\_C\_term; 1.

DR PRINTS: PRO0205; CADHERIN.

DR SMART: SM00112; CA; 5.

DR PROSITE: PS00232; CADHERIN\_1; 3.

DR PROSITE: PS02068; CADHERIN\_2; 5.

DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat; signal.

KW SIGNAL.

FT PROPEP 1 23

FT CHAIN 24 159

FT CHAIN 160 906

FT CHAIN 160 724

FT TRANSMEM 725 906

FT DOMAIN 747 946

FT DOMAIN 160 267

FT DOMAIN 268 382

FT DOMAIN 383 497

FT DOMAIN 498 603

FT DOMAIN 604 714

FT DOMAIN 863 878

FT CARBOHYD 190 190

FT CARBOHYD 273 273

FT CARBOHYD 325 325

FT CARBOHYD 402 402

FT CARBOHYD 572 572

FT CARBOHYD 651 651

FT CARBOHYD 692 692

FT CONFLICT 12 12

FT CONFLICT 16 16

FT CONFLICT 196 196

FT CONFLICT 212 212

FT CONFLICT 357 357

FT CONFLICT 867 867

SO SEQUENCE 906 AA; 99851 MW; 72DDC7B8B57C7AFC CRC64;

Query Match 42.8%; Score 38.5; DB 1; Length 906;

Best Local Similarity 50.0%; Pred. No. 43;

Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 3 PAM---YXXRGIRPGR 17

DB 414 PAMNAVRIISGDPAGR 431

Search completed: September 13, 2002, 09:30:45

Job time: 1135 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:19 ; Search time 311.85 Seconds  
(without alignments)  
9.985 Million cell updates/sec

```
Title: US-09-446-543A-73_COPY_4_21
Perfect score: 90
Sequence: 1 INPAMYXXRCIRPVGRFX 18
```

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

```
Searched:      562222 segs, 172994929 residues
Total number of hits satisfying chosen parameters: 562222
```

```

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
                  Maximum Match 100%
                  Listing first 45 summaries

```

```
Database : SPReML_arch:.*
1: sp_archaea:.*
2: sp_bacteria:.*
3: sp_fungi:.*
4: sp_human:.*
5: sp_invertebrate:.*
6: sp_mammal:.*
7: sp_mhc:.*
8: sp_organelle:.*
9: sp_phage:.*
10: sp_plant:.*
11: sp_rodent:.*
12: sp_virus:.*
13: sp_vertebrate:.*
14: sp_unclassified:.*
15: sp_virus:.*
16: sp_bacteriap:.*
17: sp_archaeap:.*
```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed and is derived by analysis of the total score distribution.

## SUMMARIES

Result	Score	Query	Match	Length	DB	ID	Description
No.							
1	72	80.0	117	13	Q9W624	Q9W624	carassius a
2	46.5	51.7	420	2	Q93L27	Q93L27	strepomyce
3	46	51.1	333	16	Q9PH76	Q9PH76	xylella fas
4	46	51.1	790	10	Q9M371	Q9M371	arabidopsis
5	44	48.9	419	4	Q9Y276	Q9Y276	homo sapien
6	43	47.8	220	16	Q9J022	Q9J022	pseudomonas
7	43	47.8	418	11	Q98ZP5	Q98ZP5	mus musculu
8	42.5	47.2	360	5	Q18879	Q18879	caenorhabdi
9	42	46.7	416	16	Q99Z49	Q99Z49	streplococc
10	42	46.7	433	16	Q9A382	Q9A382	caulobacter
11	41	45.6	97	2	Q33440	Q33440	pseudomonas
12	41	45.6	226	16	Q9A359	Q9A359	caulobacter
13	41	45.6	338	5	Q18729	Q18729	caenorhabdi
14	41	45.6	545	16	Q9A747	Q9A747	caulobacter
15	40	44.4	76	5	Q9V8L0	Q9V8L0	drosophila
16	40	44.4	153	9	Q9T133	Q9T133	bacterioph

17	40	44.4	184	2	086838	066838 streptomyces
18	40	44.4	284	17	050128	050128 streptococcus
19	40	44.4	324	2	092554	092554 streptomyces
20	40	44.4	390	16	09PM18	09PM18 xylella fast
21	40	44.4	486	16	092DM2	092DM2 listeria in
22	40	44.4	540	10	09LCS20	09LCS20 oryza sativ
23	40	44.4	562	5	09VNM4	09VNM4 drosophila
24	39.5	43.9	779	3	09PSJ9	09PSJ9 neurospora
25	39	43.3	108	11	09DLV4	09DLV4 mus musculus
26	39	43.3	179	2	09K3G7	09K3G7 streptomyces
27	39	43.3	250	16	09J1F2	09J1F2 pseudomonas
28	39	43.3	256	16	09ZVE9	09ZVE9 rhizobium m
29	39	43.3	327	10	09LUI4	09LUI4 arabidopsis
30	39	43.3	333	4	09ESD4	09ESD4 homo sapien
31	39	43.3	340	5	018731	018731 caenorhabdi
32	39	43.3	370	11	09B3V7	09B3V7 mus musculus
33	39	43.3	425	16	0968U6	0968U6 rhizobium l
34	39	43.3	555	3	000050	000050 aspergillus
35	39	43.3	938	10	09ZVE3	09ZVE3 arabidopsis
36	39	43.3	2104	5	021281	021281 caenorhabdi
37	39	43.3	2104	5	0964M4	0964M4 caenorhabdi
38	38.5	42.8	238	11	055075	055075 citreulius
39	38.5	42.8	374	6	053319	053319 mycobacteri
40	38	42.2	57	9	09ZXX0	09ZXX0 pseudomonas
41	38	42.2	127	16	09HXK7	09HXK7 streptococc
42	38	42.2	178	16	097SU7	097SU7 caulobacter
43	38	42.2	194	16	09ASU0	09ASU0 catulobacter
44	38	42.2	210	10	09LEM6	09LEM6 catharinuln
45	38	42.2	238	17	09YEH0	09YEH0 aetopyrum p

## ALIGNMENTS

[illegible]

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CC3239;  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RL Streptomyces aureofaciens CC3239.";  
DR Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
EMBL: AY033994; AAK61719.1;  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;  
  
Query Match 51.7%; Score 46.5; DB 2; Length 420;  
Best Local Similarity 52.6%; Pred. No. 4.4;  
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;  
  
OY 4 AWYXX-----RCIRPVGRF 17  
DB 37 AMMAVLGSGSIRPVGRF 55  
  
RESULT 3  
O9PH76 PRELIMINARY; PRT; 333 AA.  
AC O9PH76:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYDROXYBENZONATE OCTAPRENYLTRANSFERASE.  
GN XF0068.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OX Xylella.  
NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Rehnach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bioness M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Fetto J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitchina J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madella A.M.B.N., Madella H.M.F., Marino C.L.,  
RA Menck M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Mesquita C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira V.A. Jr., Pesquero J.B.,  
RA Quaglio R.B., Roberto P.G., Rodrigues H., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsunako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zaitz M., Meidanis J., Secubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
DR EMBL: AE003860; AAF82881.1;  
DR InterPro: IPR000537; UDLA.

DR Pfam; PF01040; UDLA.1.  
DR PROSITE; PS00943; UDLA; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 333 AA; 37931 MW; ECF3F4716C962B95 CRC64;  
  
Query Match 51.1%; Score 46; DB 16; Length 333;  
Best Local Similarity 53.3%; Pred. No. 4.2;  
Matches 8; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
  
OY 1 INPAMYXXRGIRPVG 15  
::|  
|||  
DB 54 LDPYKTLARGDRPVG 68  
  
RESULT 4  
O9M371 PRELIMINARY; PRT; 790 AA.  
AC O9M371:  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL 87.4 KDA PROTEIN.  
GN F15G16.60.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De Haan M., Maarse A.C., Grievell L.A., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Queller F., Salanoubat M.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL132959; CAB71097.1;  
KW Hypothetical protein.  
SQ SEQUENCE 790 AA; 87376 MW; B22724B75690F30 CRC64;  
  
Query Match 51.1%; Score 46; DB 10; Length 790;  
Best Local Similarity 50.0%; Pred. No. 11;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
OY 2 NPAMYXXRGIRPVGRF 17  
|||  
|||  
DB 369 NPTYSGRGLQPHGRW 384  
  
RESULT 5  
O9Y276 PRELIMINARY; PRT; 419 AA.  
AC O9Y276:  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE H-BCS1 (BCS1 (YEAST HOMOLOG)-LIKE).  
GN BCS1 OR BCS1L.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=99097350; PubMed=9878253;  
RA Petrizzella V., Tiranli V., Fernandez P., Ianna P., Carrozzo R.,  
RA Zeviani M.;  
RT "Identification and characterization of human cDNAs specific to BCS1,  
PFT12, SCO1, COX15, and COX11, five genes involved in the formation



Query Match 47.8%; Score 43; DB 11; Length 418;  
Best Local Similarity 70.0%; Pred. No. 18;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 NPAMYXXRGI 11  
|||:|:|  
Db 211 NPKWYIDRGI 220

RESULT 8  
O19879 PRELIMINARY; PRT; 360 AA.  
AC O19879;  
DT 01-NOV-1996 (TREMBLrel. 01, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE F28D1.8 PROTEIN.  
GN F28D1.8.  
OS Caenorhabditis elegans.  
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabdilitida; Rhabdilitidae;  
OC Rhabdilitidae; Pelodermine; Caenorhabditis.  
OX NCBI\_TaxID=6239;  
RN [1]  
RP SEQUENCE FROM N.A.  
RL Submitted (APR-1996) to the EMBL/GenBank/DBJ databases.  
RX [2]  
RX MEDLINE=99069613; PubMed=9851916;  
RA none;  
RT "Genome sequence of the nematode C.elegans: A platform for  
investigating biology.";  
RL Science 282:2012-2018(1996).  
DR EMBL: 270684; CAAG4603.2; -.  
SQ SEQUENCE 360 AA; 39234 MW; A5BBE127E7B3A91C CRC64;

Query Match 47.2%; Score 42.5; DB 5; Length 360;  
Best Local Similarity 37.5%; Pred. No. 19;  
Matches 9; Conservative 3; Mismatches 3; Indels 9; Gaps 1;

OY 1 INPAMYXXR-----GIRVVG 15  
|||:|:|  
Db 53 INPSWEDRWVSYSIDGHLGIMPVG 76

RESULT 9  
O092A9 PRELIMINARY; PRT; 416 AA.  
AC O092A9;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PUTATIVE EXTRAMEMBRANAL PROTEIN.  
GN D1TD OR SPY1309.  
OS Streptococcus pyogenes.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1314;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=SF370 / ATCC 700294 / SEROTYPE M1;  
RX MEDLINE=21192684; PubMed=11296296;  
RA Ferrerelli J.J., Moshan W.M., Ajdic D.J., Savic D.J., Savic G., Lyon K.,  
Primeaux C., Sezate S., Suvorov A.N., Kenton S., Lai H.S., Lin S.P.,  
Qian Y., Jia H.G., Najjar F.Z., Ren Q., Zhu H., Song L., White J.,  
Yuan X., Clifton S.W., Roe B.A., McLaughlin R.;  
RT "Complete genome sequence of an M1 strain of Streptococcus pyogenes.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4658-4663(2001).  
DR EMBL: AE006570; AAK34153.1; -.  
KW Complete proteome.

SQ SEQUENCE 416 AA; 48068 MW; EACAA2203F7E519F CRC64;

Query Match 46.7%; Score 42; DB 16; Length 416;  
Best Local Similarity 38.5%; Pred. No. 27;  
Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 INPAMYXXRGIR 13  
|||:|:|  
Db 129 VSPQWFTAGINP 141

RESULT 10  
O0A382 PRELIMINARY; PRT; 433 AA.  
ID O0A382  
AC O0A382;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PEPTIDOLYCAN-BINDING PROTEIN, PUTATIVE.  
GN CC3322.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=ATCC 19089 / CB15;  
RX MEDLINE=21173698; PubMed=11259647;  
RA Nierman W.C., Feldblyum T.V., Laud M.T., Paulsen I.T., Nelson K.E.,  
Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
Rapoport J., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
DeBoy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,  
Kolony J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,  
Ruttenberg S.L., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
RT "Complete genome sequence of Caulobacter crescentus.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
DR EMBL: AE005994; AAK25284.1; -.  
DR HSSP; P41052; ILTM.  
DR TIGR; CC3322; -.  
DR InterPro: IPR002477; PG\_binding.  
DR Pfam; PF01471; PG\_binding\_1; 1.  
KW Complete proteome.  
SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

Query Match 46.7%; Score 42; DB 16; Length 433;  
Best Local Similarity 60.0%; Pred. No. 29;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 3 PAMYXXRGIR 12  
|||:|:|  
Db 285 PAMWEARQVR 294

RESULT 11  
O03440 PRELIMINARY; PRT; 97 AA.  
ID O03440  
AC O03440;  
DT 01-JAN-1998 (TREMBLrel. 05, Created)  
DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE NADH-UBIQUINONE OXIDOREDUCTASE (FRAGMENT).  
GN NUO4.  
OS Pseudomonas fluorescens.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=294;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX STRAIN=WCS365;  
RX MEDLINE=96340543; PubMed=9675892;



RA Dekkers J.C., van der Bij A.J., Mulders H.M., Pheollet C.C.,  
RA Wentwood A.R., Glandorf D.C.M., Wijffelman C.A., Lugtenberg B.J.J.,  
RT "Role of the O-antigen of lipopolysaccharide, and possible roles of  
RT growth rate and NADH:ubiquinone oxidoreductase (nuo) in competitive  
RT tomato root-tip colonization by *Pseudomonas fluorescens* WCS365.";  
RL Mol. Plant Microbe Interact. 11:763-771(1998).  
DR EMBL: Y14569; CAA74900.1; -  
DR InterPro: IPR001135; Complex1\_49Kd.  
DR Pfam: PF00346; Complex1\_49Kd; 1.  
FW Ubiquinone; NAD.  
FW Ubiquinone; NAD.  
FT NON\_TER 1 1  
FT NON\_TER 97 97  
SO SEQUENCE 97 AA; 11450 MW; 6568C662A3939D9 CRC64;

Q9V8L0  
 ID Q9V8L0 PRELIMINARY; PRT; 76 AA.  
 AC Q9V8L0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last annotation update)  
 DE CG12613 PROTEIN.  
 GN CG12613.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OK NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 Amaratilake P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 Brandon R.C., Rogers Y.-H.C., Blazer V., Chapple M., Pfeiffer B.D.,  
 Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 Durbin K.J., Evangelista C.C., Ferrara C., Fertl S., Fleischmann W.,  
 Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 Laako P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
 Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclio J.M.,  
 Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 Splet E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 Svirskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissenbach J.,  
 Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Glbbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL; AE003798; AAF57654.1;-  
 DR FlyBase; FBgn0034375; CG12613.  
 SO SEQUENCE 76 AA; 8646 MW; 0BBE8F08E1DB70ED CRC64;

Query Match 44.4%; Score 40; DB 5; Length 76;  
 Best Local Similarity 54.5%; Pred. NO. 9.8;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 INPARYXXRGI 11  
 :||| |  
 Db 25 VNPGRTEAREI 35

Search completed: September 13, 2002, 09:29:20  
 Job time: 1065 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OW protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds  
(without alignments)  
4.724 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_5\_21  
Perfect score: 86  
Sequence: 1 NPAMWXXRCIRPVGRFX 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

- 1: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	96.5	20	18 AAW31394	Human type G prote
2	83	96.5	20	20 AAW97236	Human type ligand
3	83	96.5	20	21 AAB10365	Human oxytocin sec
4	83	96.5	20	21 AAY49294	19P2 ligand peptid
5	83	96.5	20	22 AAG62534	Human CRH releasin
6	83	96.5	20	22 AAB90992	Prolactin releasin
7	83	96.5	21	18 AAW31395	Human type G prote
8	83	96.5	21	21 AAB10366	Human oxytocin sec
9	83	96.5	21	22 AAG62535	Human CRH releasin
10	83	96.5	22	18 AAW31396	Human type G prote
11	83	96.5	22	21 AAB10367	Human oxytocin sec

12	83	96.5	22	22 AAG62536	Human CRH releasin
13	83	96.5	31	18 AAW31391	Human type G prote
14	83	96.5	31	20 AAW97235	Human type ligand
15	83	96.5	31	20 AAB87615	Human 19P2 ligand
16	83	96.5	31	21 AAB10362	Human oxytocin sec
17	83	96.5	31	21 AAY49291	19P2 ligand peptid
18	83	96.5	31	22 AAG62531	Human CRH releasin
19	83	96.5	31	22 AAB90991	Prolactin releasin
20	83	96.5	31	22 AAB90995	Prolactin releasin
21	83	96.5	32	18 AAB31392	Human type G prote
22	83	96.5	32	21 AAB10363	Human oxytocin sec
23	83	96.5	32	22 AAG62532	Human CRH releasin
24	83	96.5	33	18 AAW31393	Human type G prote
25	83	96.5	33	21 AAB10364	Human oxytocin sec
26	83	96.5	33	22 AAG62533	Human CRH releasin
27	83	96.5	87	18 AAW31390	Human type G prote
28	83	96.5	87	20 AAW97226	Human type ligand
29	83	96.5	87	21 AAB10361	Human oxytocin sec
30	83	96.5	87	22 AAG62530	Human CRH releasin
31	82	95.3	20	18 AAW31387	Rat type G protein
32	82	95.3	20	18 AAW31374	Bovine G protein-c
33	82	95.3	20	20 AAW97232	Bovine pituitary-d
34	82	95.3	20	20 AAW97234	Rat type ligand po
35	82	95.3	20	20 AAW95191	Bovine pituitary-d
36	82	95.3	20	21 AAW95175	Murine pituitary-d
37	82	95.3	20	21 AAB10350	Bovine oxytocin se
38	82	95.3	20	21 AAB10358	19P2 ligand peptid
39	82	95.3	20	21 AAY49301	19P2 ligand peptid
40	82	95.3	20	21 AAY49302	Bovine CRH releasi
41	82	95.3	20	22 AAG62519	19P2 ligand peptid
42	82	95.3	20	22 AAG62527	Rat CRH releasin
43	82	95.3	20	22 AAB90994	Prolactin releasin
44	82	95.3	20	22 AAB90996	Prolactin releasin
45	82	95.3	20	22 AAB46954	peptide PRP20 fra

#### ALIGNMENTS

RESULT 1

AAW31394 standard; Peptide: 20 AA.

AAW31394;

06-APR-1998 (first entry)

Human type G protein-coupled receptor ligand fragment 4.

G protein-coupled receptor; ligand binding; pharmaceutical;

modulator; pituitary; central nervous system; pancreas; prophylactic;

therapeutic agent.

Homo sapiens.

WO9724436-A2.

10-JUL-1997.

26-DEC-1996; 96WO-JP03821.

18-SEP-1996; 96JP-0246573.

28-DEC-1995; 95JP-0343371.

15-MAR-1996; 96JP-0059419.

12-AUG-1996; 96JP-0211805.

(TAKE ) TAKEDA CHEM IND LTD.

Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

Kawamata Y, Kitada C;

WPI: 1997-363672/33.

N-PSDB; AAV02431.

XX Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 XX Claim 2; Page 185; 258pp; English.  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
 CC sequence represented in AAM97236 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 CC  
 XX  
 SQ Sequence 20 AA;

Query Match 96.5%: Score 83; DB 18; Length 20;  
 Best Local Similarity 87.5%: Pred. No. 3.6e-08;  
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
 ||||| |||||  
 Db 5 npawyaargirpvgrf 20

RESULT 2  
 AAM97236  
 ID AAM97236 standard; peptide; 20 AA.  
 XX  
 AC AAM97236;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

XX  
 KW Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopause; syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
 KW acromegaly; Charli-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
 KW interruption mole; abortion; unfertility fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 XX

OS Homo sapiens.

PN WO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 XX

DR WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopause  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 XX Claim 3; Page 166; 241pp; English.  
 XX

CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyst cacogenesis, menopause  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenorrhoea, galactorrhea,  
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
 CC acromegaly, Charli-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, interruption mole, abortion, unfertility fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 CC  
 XX  
 SQ Sequence 20 AA;

Query Match 96.5%: Score 83; DB 20; Length 20;  
 Best Local Similarity 87.5%: Pred. No. 3.6e-08;  
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
 ||||| |||||  
 Db 5 npawyaargirpvgrf 20

RESULT 3  
 AAB10365  
 ID AAB10365 standard; peptide; 20 AA.  
 XX  
 AC AAB10365;

DT 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.

XX  
 KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX

OS Homo sapiens.

PN WO200038704-A1.

PD 06-JUL-2000.

PF 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -

XX Disclosure: Page 63; 72pp; Japanese.

PS  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.

CC  
XX  
SQ Sequence 20 AA;

Query Match 96.5%; Score 83; DB 21; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.6e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
||||| |||||||||  
DB 5 npawyasrigrpvgrf 20

RESULT 4  
AAV49294  
ID AAV49294 standard; peptide: 20 AA.  
AC AAV49294;  
XX  
XX 22-FEB-2000 (first entry)  
DT  
XX  
DE 19P2 ligand peptide fragment.  
XX  
XX Monoclonal antibody; 19P2 ligand; diagnosis: prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
KM  
XX Homo sapiens.  
OS  
XX  
FH Key Location/Qualifiers  
FT Modified-site 20 /note="C-terminal amide"  
ET  
XX  
XX WO9960112-A1.  
PM  
XX  
XX 25-NOV-1999.  
PD  
XX  
XX 20-MAY-1999; 99WO-JP02650.  
PE  
XX  
XX 21-MAY-1998; 98JP-0140293.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
PI  
XX  
XX WPI; 2000-039381/03.  
DR  
XX  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality  
PS  
XX  
XX Disclosure: Page 26; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunosay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAV49290-302 represent peptide fragments of the 19P2 ligand.

XX  
SQ Sequence 20 AA;

Query Match 96.5%; Score 83; DB 21; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.6e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
||||| |||||||||  
DB 5 npawyasrigrpvgrf 20

RESULT 5  
AAG62534  
ID AAG62534 standard; peptide: 20 AA.  
AC AAG62534;  
XX  
XX 24-AUG-2001 (first entry)  
DT  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 35.  
XX  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
KM  
XX Homo sapiens.  
OS  
XX  
XX WO200135984-A1.  
PN  
XX  
XX 25-MAY-2001.  
PD  
XX  
XX 17-NOV-2000; 2000WO-JP08119.  
PE  
XX  
XX 18-NOV-1999; 99JP-0327900.  
PR  
XX  
XX 26-SEP-2000; 2000JP-0297073.  
PA  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PI  
XX  
XX Kitada C, Matsumoto H, Hinuma S;  
XX  
XX WPI; 2001-355552/37.  
DR  
XX  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
PS  
XX  
XX Claim 4; Page 75; 90pp; Japanese.

CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.

CC  
XX  
SQ Sequence 20 AA;

Query Match 96.5%; Score 83; DB 22; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.6e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
||||| |||||||||  
DB 5 npawyasrigrpvgrf 20

RESULT 6

AAB90992  
ID AAB90992 standard; Peptide: 20 AA.  
XX  
AC AAB90992;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Prolactin releasing peptide SEQ ID NO:166.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
PA (CONU-) CONUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI: 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PT  
XX  
PS Disclosure; Page 244; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specifically as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 20 AA;  
XX  
Query Match 96.5%; Score 83; DB 22; Length 20;  
Best Local Similarity 87.5%; Pred. NO. 3.6e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXXKGRPVGRF 16  
Db 5 npawyxxgrlrvpgrf 20

RESULT 7  
AAB90992  
ID AAB90992 standard; Peptide: 21 AA.  
XX  
AC AAB90992;

XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 5.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prolactin;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR N-PSDB; AAV02432.  
XX  
DR WPI: 1997-363672/33.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
PT  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
XX This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AAB90992 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC Turner's syndrome, diabetes, cancer, pancreatitis, renal disease,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosacchara. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 21 AA;  
XX  
Query Match 96.5%; Score 83; DB 18; Length 21;  
Best Local Similarity 87.5%; Pred. NO. 3.8e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXXKGRPVGRF 16  
Db 5 npawyxxgrlrvpgrf 20

RESULT 8  
AAB10366  
ID AAB10366 standard; Peptide: 21 AA.  
XX  
AC AAB10366;

DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI: 2000-452298/39.  
XX  
XX Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
PS Disclosure: Page 63; 72pp; Japanese.  
XX  
XX This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
CC  
XX  
SQ Sequence 21 AA:  
  
Query Match 96.5%; Score 83; DB 21; Length 21;  
Best Local Similarity 87.5%; Pred. No. 3.8e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 NPAMYXXRGIRPYGRF 16  
| | | | | | | | | | | | | | | | | | | | | |  
DB 5 npawyasrglrpygrf 20  
  
RESULT 9  
AAC62535  
ID AAC62535 standard; peptide; 21 AA.  
XX  
AC AAG62535;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 36.  
XX  
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX

PD 25-MAY-2001.  
XX  
XX 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
XX  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
XX  
DR WPI: 2001-355552/37.  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Disclosure: Page 75; 90pp; Japanese.  
XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 21 AA:  
  
Query Match 96.5%; Score 83; DB 22; Length 21;  
Best Local Similarity 87.5%; Pred. No. 3.8e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 NPAMYXXRGIRPYGRF 16  
| | | | | | | | | | | | | | | | | | | | | |  
DB 5 npawyasrglrpygrf 20  
  
RESULT 10  
AAW31396  
ID AAW31396 standard; Peptide; 22 AA.  
XX  
AC AAW31396;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 6.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
XX  
PR 28-DEC-1995; 95JP-0343771.  
XX  
PR 15-MAR-1996; 96JP-0059419.  
XX  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX

DR WPI: 1997-363672/33.  
DR N-PSDB; AAV02433.  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2; Page 186; 258pp; English.  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC truma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharitis. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
XX  
SQ Sequence 22 AA:  
  
Query Match 96.5%; Score 83; DB 18; Length 22;  
Best Local Similarity 87.5%; Pred. No. 3.9e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 NPAWYXXRGIRPVGRF 16  
| | | | | | | | | | | | | | | | | |  
DB 5 npawyasrgirpvgrf 20  
  
RESULT 11  
ID AAB10367 standard; peptide: 22 AA.  
XX  
XX AAB10367;  
XX  
XX 24-NOV-2000 (first entry)  
XX  
XX Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
DE  
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
XX Homo sapiens.  
XX  
XX WO200038704-A1.  
XX  
XX PN  
XX PD 06-JUL-2000.  
XX  
XX PF 22-DEC-1999; 99WO-JP07199.  
XX  
XX PR 25-DEC-1998; 98JP-0369585.  
XX  
XX PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX PI Matsumoto H, Kitada C, Hinuma S;  
XX  
XX WPI: 2000-452298/39.  
XX  
XX DR  
XX PT Physiologically-active polypeptide recognized as ligand by G  
XX protein-coupled receptor protein, for promoting secretion of oxytocin,

PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
XX  
PS Disclosure; Page 64; 72pp; Japanese.  
XX  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
XX  
SQ Sequence 22 AA:  
  
Query Match 96.5%; Score 83; DB 21; Length 22;  
Best Local Similarity 87.5%; Pred. No. 3.9e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 NPAWYXXRGIRPVGRF 16  
| | | | | | | | | | | | | | | | | |  
DB 5 npawyasrgirpvgrf 20  
  
RESULT 12  
ID AAG62536 standard; peptide: 22 AA.  
XX  
XX AAG62536;  
XX  
XX 24-NOV-2001 (first entry)  
XX  
XX DE Human CRH releasing protein related peptide SEQ ID NO: 37.  
XX  
XX KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analogs; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX OS Homo sapiens.  
XX  
XX PN WO200135984-A1.  
XX  
XX PD 25-MAY-2001.  
XX  
XX PF 17-NOV-2000; 2000WO-JP08119.  
XX  
XX PR 18-NOV-1999; 99JP-0327900.  
XX  
XX PR 26-SEP-2000; 2000JP-0297073.  
XX  
XX PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX PI Kitada C, Matsumoto H, Hinuma S;  
XX  
XX WPI: 2001-35552/37.  
XX  
XX DR  
XX PF Use of G protein receptor ligand or peptide for controlling  
XX corticotropin releasing hormone secretion -  
XX  
XX PS Disclosure; Page 75; 90pp; Japanese.  
XX  
XX CC The present sequence describes a method of controlling the secretion of  
XX corticotrophin releasing hormone (CRH), involving the use of a G protein  
XX receptor ligand. This can be used to control the secretion of CRH and is  
XX useful as an analgesic or for treating, preventing or ameliorating  
XX diseases associated with CRH secretion such as hyperaldosteronism,  
XX hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
XX disease (including borden, nausea, pigmentation, hypogonadism, hair  
XX loss, and hypotension), adrenal gland hypofunction and obesity. The  
XX present sequence is a peptide used in the exemplification of the



CC Invention.  
XX  
SQ Sequence 22 AA;

Query Match 96.5%; Score 83; DB 22; Length 22;  
Best Local Similarity 87.5%; Pred. No. 3.9e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMWXXRGIRPYGRF 16  
||||| |||||||  
DB 5 NPAMWASRGIRPYGRF 20

RESULT 13  
AAW31391  
ID AAW31391 standard; Peptide; 31 AA.  
XX  
AC AAW31391;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 1.

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prolactin;  
KW therapeutic agent.

OS Homo sapiens.

XX W09724436-A2.

PM 10-JUL-1997.

PD 26-DEC-1996; 96WO-JP03821.

PF 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-034371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX WPI: 1997-363672/33.

DR N-PSDB; AAW02428.

XX Claim 2; Page 184; 258pp; English.

This sequence represents a peptide fragment from a novel human type  
ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
sequence represented in AAW31390 and is used in an assay to monitor  
ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
compositions containing this ligand may be used as a pituitary function  
modulator, a central nervous system modulator or a pancreatic function  
modulator. This ligand could have specific applications as a  
prolactin, growth hormone secretory disease, hyper- and polyphagia,  
trauma, growth hormone secretory disease, hyper- and polyphagia,  
hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
acute myocardial infarction, infertility, spinocerebellar degeneration,  
bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
oligodactylia. Assays can also be developed to screen compounds which are  
capable of altering the binding activity of the ligand affecting  
activation of the G protein-coupled receptor protein.

XX  
SQ Sequence 31 AA;

Query Match 96.5%; Score 83; DB 18; Length 31;  
Best Local Similarity 87.5%; Pred. No. 5.7e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMWXXRGIRPYGRF 16  
||||| |||||||  
DB 16 NPAMWASRGIRPYGRF 31

RESULT 14  
AAW97235  
ID AAW97235 standard; Peptide; 31 AA.  
XX  
AC AAW97235;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Human type ligand polypeptide fragment.

KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; amenorrhoea; autoimmune disease;  
KW acromegaly; infertility; impotence; amenorrhoea; galactorrhea;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
KW interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.

OS Homo sapiens.

XX W09658962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PR (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

PI WPI: 1999-105614/09.

Use of G protein-coupled receptor ligands - for modulating prolactin  
secretion or placental function, e.g. for treating menopausal  
syndrome, tumours, autoimmune disease or abnormal pregnancy

Claim 3; Page 159; 241pp; English.

The present sequence represents a human type ligand fragment. It  
is used in the course of the invention. The specification describes  
an agent for modulating prolactin secretion which comprises a  
ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
protein. The agents for promoting prolactin secretion can be used for  
treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal  
syndrome, euthyroid or hypometabolism. They can be used for promoting  
lactation in a domestic mammal and as an aphrodisiac. The agents for  
inhibiting prolactin secretion can be used for treating or preventing  
pituitary adenomatosis, brain tumour, emmenorrhoea, autoimmune disease,  
prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
acromegaly, Chlari-Frommel syndrome, Argon-z-del Castillo syndrome,  
Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
The inhibitory agents can also be used as contraceptives. The agents for  
modulating placental function can be used for treating or preventing  
choriocarcinoma, hydralid mole, abortion, unfertilized fetus,  
abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX	Sequence	31	AA;
5Q			

Query Match	96.5%	Score 83:	DB 20:	Length 31:
Best Local Similarity	87.5%	Pred. No.	5.7e-08:	
Matches 14, Conservative	0:	Mismatches 2:	Indels 0:	Gaps 0:

```
Qy 1 NPAMYXXRGIRPVGRE 16
    ||||| |||||
Db 16 NPAYASRGLRPVGRF 31
```

RESULT	15
AAW87615	
ID	AAW87615 standard; Peptide; 31 AA

AC AAW87615;

DT 29-MAR-1999 (first entry)  
yy

Human 19P2 ligand.

KW 19p2 11yand; G protein coupled receptor; pituitary;  
 KW prolactin releasing peptide; human; dementia; breast cancer;  
 KW therapy.

**Homo sapiens.**

PN EP887417-A2.

PD 30-DEC-1998.

PF 25-JUN-1998; 98EP-0111725.

PR 27-JUN-1997; 97JP-0172118.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Moriya T, Nishimura O, Suenaga M, Tanaka Y;

DR WPT; 1999-047884/05.

PT Producing a 19p2 pituitary G protein receptor ligand - by cleavage  
PT of a fusion protein, useful for preventing and treating dementia,  
PT breast cancer, renal failure and autoimmune disease

PS Claim 5; Page 35; 56pp; English.

This is the amino acid sequence of the human pituitary G protein-coupled receptor ligand 19pL2L. A method suitable for commercial high-level production of 19pL2L comprises expressing the ligand in host cells as a recombinant fusion protein e.g. with human basic fibroblast growth factor (see AAV83736-97) that has been modified to include an N-terminal cysteine residue. The ligand is released from the fusion by cyanatation followed by ammonolysis. 19pL2L has prolactin secretion-stimulating and (at high doses) prolactin secretion-inhibiting properties. It can be used in the treatment and prevention of various diseases including: senile dementia, cerebrovascular dementia, and dementia associated with: neurological disorders (e.g. Alzheimer's disease, Parkinson's disease, Pick's disease, Huntington's disease), infectious diseases (e.g. Creutzfeldt-Jakob's), endocrine or metabolic disease or toxicosis (e.g. hypothyroidism, vitamin B12 deficiency, alcoholism, intoxication by drugs, metal and organic compounds), tumorigenic diseases (e.g. brain tumour), traumatic diseases (e.g. chronic subarachnoidal hemorrhage), and other types of dementia, depression hyperactive child syndrome (microencephalopathy) and disturbance of consciousness. It is also useful for prevention and treatment of diseases associated with prolactin hypo and hypersecretion respectively, including: hyperprolactinemia, pituitary adenoma, breast cancer, infertility, impotence and autoimmune disease (hyperscretion disorders), anesmenal vesicle hypoplasia,

CC osteoporosis, menopausal syndrome and renal failure (hyposecretion  
CC disorders). The 19p2 polypeptide/amide is also useful as a test  
CC reagent for study of the prolactin secretory function or as a  
CC lactagogue in mammalian farm animals.

**SQ Sequence 31 AA;**

Query Match	96.5%;	Score 83;	DB 20;	Length 31;
Best Local Similarity	87.5%;	Pred. No. 5.7e-08;		
Matches 14;	Conservative 0;	Mismatches 2;	Indels 0;	Gaps 0;

```
OY 1 NPAWYXXRGIRPVGRF 16
    ||||| |||||
Db 16 NPAWYASRGIRPVGRF 31
```

Search completed: September 13, 2002, 09:18:35  
Job time: 500 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 Seconds  
(without alignments)  
3.039 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_5\_21  
Perfect score: 86  
Sequence: 1 NPANXXXXGIRPVGFRFX 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 segs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep: \*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep: \*  
3: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep: \*  
4: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep: \*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep: \*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	96.5	20	3	US-09-105-678A-46
2	83	96.5	20	4	US-08-776-971-64
3	83	96.5	20	4	US-09-421-208-46
4	83	96.5	21	3	US-09-105-678A-47
5	83	96.5	21	4	US-08-776-971-65
6	83	96.5	21	4	US-09-421-208-47
7	83	96.5	22	3	US-09-105-678A-48
8	83	96.5	22	4	US-08-776-971-66
9	83	96.5	22	4	US-09-421-208-48
10	83	96.5	31	3	US-09-105-678A-9
11	83	96.5	31	3	US-09-105-678A-43
12	83	96.5	31	4	US-08-776-971-61
13	83	96.5	31	4	US-09-421-208-9
14	83	96.5	31	4	US-09-421-208-43
15	83	96.5	32	3	US-09-105-678A-44
16	83	96.5	32	4	US-08-776-971-62
17	83	96.5	32	4	US-09-421-208-44
18	83	96.5	33	3	US-09-105-678A-45
19	83	96.5	33	4	US-08-776-971-63
20	83	96.5	33	4	US-09-421-208-45
21	83	96.5	87	4	US-08-776-971-59
22	83	96.5	87	4	US-08-776-971-135
23	83	96.5	87	4	US-08-776-971-138
24	83	96.5	20	3	US-09-105-678A-34
25	82	95.3	20	3	US-09-105-678A-40
26	82	95.3	20	4	US-08-776-971-8
27	82	95.3	20	4	US-08-776-971-50

28	82	95.3	20	4	US-08-776-971-98	Sequence 98, Appl
29	82	95.3	20	4	US-09-421-208-34	Sequence 34, Appl
30	82	95.3	20	4	US-09-421-208-40	Sequence 40, Appl
31	82	95.3	21	3	US-09-105-678A-35	Sequence 35, Appl
32	82	95.3	21	3	US-09-105-678A-41	Sequence 41, Appl
33	82	95.3	21	4	US-08-776-971-9	Sequence 9, Appl1
34	82	95.3	21	4	US-08-776-971-51	Sequence 51, Appl
35	82	95.3	21	4	US-09-421-208-35	Sequence 35, Appl
36	82	95.3	21	4	US-09-421-208-41	Sequence 41, Appl
37	82	95.3	22	3	US-09-105-678A-36	Sequence 36, Appl
38	82	95.3	22	3	US-09-105-678A-42	Sequence 42, Appl
39	82	95.3	22	4	US-08-776-971-10	Sequence 10, Appl
40	82	95.3	22	4	US-08-776-971-52	Sequence 52, Appl
41	82	95.3	22	4	US-09-421-208-36	Sequence 36, Appl
42	82	95.3	22	4	US-09-421-208-42	Sequence 42, Appl
43	82	95.3	31	3	US-09-105-678A-7	Sequence 7, Appl1
44	82	95.3	31	3	US-09-105-678A-8	Sequence 8, Appl1
45	82	95.3	31	3	US-09-105-678A-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
; Sequence 46, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Motiyya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-46

Query Match 96.5%; Score 83; DB 3; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAWYXRGIRPVGRF 16  
||||| |||||||  
Db 5 NPAWYASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64  
; Sequence 64, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 96.5%; Score 83; DB 4; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAWYXRGIRPVGRF 16  
||||| |||||||  
Db 5 NPAWYASRGIRPVGRF 20

RESULT 3  
US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3440  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 96.5%; Score 83; DB 4; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAWYXRGIRPVGRF 16  
||||| |||||||  
Db 5 NPAWYASRGIRPVGRF 20

RESULT 4  
US-09-105-678A-47  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston

Query Match 96.5%; Score 83; DB 4; Length 20;  
Best Local Similarity 87.5%; Pred. No. 3.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 96.5%; Score 83; DB 3; Length 21;  
Best Local Similarity 87.5%; Pred. No. 3.3e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAWYXXRGIRPYGRF 16  
||||| |||||||  
DB 5 NPAWYASRGIRPYGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hironuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: PastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 96.5%; Score 83; DB 4; Length 21;  
Best Local Similarity 87.5%; Pred. No. 3.3e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAWYXXRGIRPYGRF 16  
||||| |||||||  
DB 5 NPAWYASRGIRPYGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 96.5%; Score 83; DB 4; Length 21;  
Best Local Similarity 87.5%; Pred. No. 3.3e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
||||| |||||||  
Db 5 NPAMYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 96.5%; Score 83; DB 3; Length 22;  
Best Local Similarity 87.5%; Pred. No. 3.3e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
||||| |||||||  
Db 5 NPAMYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawabata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 96.5%; Score 83; DB 4; Length 22;  
Best Local Similarity 87.5%; Pred. No. 3.5e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
||||| |||||||  
Db 5 NPAMYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

; Sequence 48, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-48

Query Match 96.5%; Score 83; DB 4; Length 22;  
Best Local Similarity 87.5%; Pred. No. 3.5e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXRGIRPVGRF 16  
||||| |||||||  
DB 5 NPAMYASRGIRPVGRF 20

RESULT 10  
US-09-105-678A-9  
; Sequence 9, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109

; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-9

Query Match 96.5%; Score 83; DB 3; Length 31;  
Best Local Similarity 87.5%; Pred. No. 5.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXRGIRPVGRF 16  
||||| |||||||  
DB 16 NPAMYASRGIRPVGRF 31

RESULT 11  
US-09-105-678A-43  
; Sequence 43, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440

;; INFORMATION FOR SEQ ID NO: 43:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 31 amino acids  
;; TYPE: amino acid  
;; STRANDEDNESS:  
;; TOPOLOGY: linear  
US-09-105-678a-43

Query Match 96.5%; Score 83; DB 3; Length 31;  
Best Local Similarity 87.5%; Pred. No. 5.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXRGIRPVGRF 16  
Db 16 NPAMYASRGIRPVGRF 31  
||||| |||||||

RESULT 12  
US-08-776-971-61  
; Sequence 61, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 61:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear

;; MOLECULE TYPE: protein  
;; FRAGMENT TYPE: internal  
;; SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 96.5%; Score 83; DB 4; Length 31;  
Best Local Similarity 87.5%; Pred. No. 5.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXRGIRPVGRF 16  
Db 16 NPAMYASRGIRPVGRF 31  
||||| |||||||

RESULT 13  
US-09-421-208-9  
; Sequence 9, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 9:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 96.5%; Score 83; DB 4; Length 31;  
Best Local Similarity 87.5%; Pred. No. 5.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAMYXRGIRPVGRF 16  
Db 16 NPAMYASRGIRPVGRF 31  
||||| |||||||

RESULT 14



US-09-421-208-43  
; Sequence 43, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27, 026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-43

Query Match 96.5%; Score 83; DB 4; Length 31;  
Best Local Similarity 87.5%; Pred. No. 5.1e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAWYXXRGIRPVGRF 16  
DB 16 NPAWYASRGIRPVGRF 31

RESULT 15  
US-09-105-678A-44  
; Sequence 44, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27, 026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-44

Query Match 96.5%; Score 83; DB 3; Length 32;  
Best Local Similarity 87.5%; Pred. No. 5.3e-08;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAWYXXRGIRPVGRF 16  
DB 16 NPAWYASRGIRPVGRF 31

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

THIS PAGE BLANK (USPTO)

THIS PAGE BLANK (USPTO)

11/11/11

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

# OM protein - protein search, using sw model

Run on: September 13, 2002, 09:23:59 ; Search time 172.41 Seconds  
(without alignments)  
9.475 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_5\_21  
Perfect score: 86  
Sequence: 1 NPAWYXXRGIRPYGRFX 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	82	95.3	83	2 JC7607	prolactin-releasing
2	46	53.5	790	2 T47959	hypothetical prote
3	44	51.2	333	2 H82852	hydroxybenzoate oc
4	43	50.0	220	2 C83292	probable glutathio
5	43	50.0	719	2 S61046	ARF1 protein - yea
6	42	48.8	433	2 H87660	peptidoglycan-bind
7	41	47.7	338	2 T20100	hypothetical prote
8	41	47.7	476	2 G64720	probable amino aci
9	41	47.7	476	2 G90629	probable inner mem
10	41	47.7	476	2 G85480	inner membrane tra
11	41	47.7	476	2 AG0502	probable amino-act
12	41	47.7	545	2 AB7448	conserved hypothet
13	40	46.5	184	2 T35841	probable membrane
14	40	46.5	324	2 T35901	probable arac faml
15	40	46.5	390	2 G82844	cysteine synthase
16	40	46.5	486	2 AF1174	lysine-specific pe
17	40	46.5	486	2 AG1531	lysine-specific pe
18	40	46.5	965	2 AG1531	valine--trRNA ligas
19	39	45.3	149	1 NEBP37	endodeoxyribonuc
20	39	45.3	256	2 E95936	hypothetical prote
21	39	45.3	340	2 T20102	hypothetical prote
22	39	45.3	485	2 E83663	glutamyl-CRNA synt
23	39	45.3	938	2 E84480	valyl-CRNA syntet
24	39	45.3	953	2 E82068	hypothetical prote
25	38.5	44.8	443	2 T21499	N-cadherin precurs
26	38.5	44.8	877	1 IUBOCN	N-cadherin 2 precurs
27	38.5	44.8	906	1 IJHUCN	N-cadherin precurs
28	38.5	44.8	906	1 IJHUCN	N-cadherin precurs
29	38	44.2	127	2 B83157	hypothetical prote

30	38	44.2	158	2 F72725	probable ribosomal
31	38	44.2	167	2 AB2796	acetyltransferase
32	38	44.2	167	2 B97575	hypothetical prote
33	38	44.2	178	2 D95026	ribosomal protein
34	38	44.2	178	2 D97897	50S ribosomal prot
35	38	44.2	194	2 D87357	hypothetical prote
36	38	44.2	226	2 H72646	hypothetical prote
37	38	44.2	238	2 H72646	hypothetical prote
38	38	44.2	342	2 B64395	malic acid transpo
39	38	44.2	347	2 H64371	oxidoreductase alp
40	38	44.2	359	2 T35179	acetate kinase - D
41	38	44.2	398	2 B75254	valyl-CRNA syntet
42	38	44.2	950	2 G83167	valine--trRNA ligas
43	38	44.2	951	1 SYECVT	valine trRNA synthe
44	38	44.2	951	2 C91283	valine trRNA synthe
45	38	44.2	951	2 E86124	valine trRNA synthe

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishioka, T.; Hosoya, Blochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRRP  
A:Introns: 33/1

Query Match 95.3%; Score 82; DB 2; Length 83;  
Best Local Similarity 87.5%; Pred. No. 2.6e-07;  
Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 NPAWYXXRGIRPYGRF 16  
|||||  
Db 37 NPAWYXXRGIRPYGRF 52

RESULT 2  
T47959  
hypothetical protein F15G16.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Nemes, H.W.; Lemcke, K.; Mayer, K.F.X.; Q submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z24480  
A:Accession: T47959  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEH>  
A:Cross-references: EMBL:AL12959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:Introns: 39/1; 678/2; 698/3; 773/2  
A:Note: F15G16.60



A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-433 <STO>  
A:Cross-references: GB:AE005673; NID:g13425020; PIDN:AAK25284.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3322

Query Match 48.8%; Score 42; DB 2; Length 433;  
Best Local Similarity 60.0%; Pred. No. 14;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAMYXXRGIR 11  
||| ||:|  
Db 285 PAMWEARGVR 294

## RESULT 7

hypothetical protein G50C10.2 - *Caenorhabditis elegans*  
C:Species: *Caenorhabditis elegans*  
C>Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T20100  
R:McMurray, A.  
submitted to the EMBL Data Library, May 1996  
A:Reference number: Z19224  
A:Accession: T20100  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-338 <WIL>  
A:Cross-references: EMBL:Z72505; PIDN:CAA96608.1; GSPDB:GN00023; CESP:C50C10.2  
A:Experimental source: clone C50C10  
C:Genetics:  
A:Gene: CESP:C50C10.2  
A:Map position: 5  
A:introns: 74/3; 144/3; 267/3

Query Match 47.7%; Score 41; DB 2; Length 338;  
Best Local Similarity 53.8%; Pred. No. 16;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 2 PAMYXXRGIRPVG 14  
||| ||:|  
Db 275 PYWYQLEIRPIG 287

## RESULT 8

probable amino acid transport protein yaaU, sodium-dependent - *Escherichia coli*  
C:Species: *Escherichia coli*  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
C:Accession: G64720  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G64720  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <BLAT>  
A:Cross-references: GB:AE000111; GB:U00096; NID:g1786181; PIDN:ACG73118.1; PID:g1786188;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yaaU  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein  
C:Keywords: amino acid transport; transmembrane protein  
F:10-26/Domain: transmembrane #status predicted <TM1>  
F:91-107/Domain: transmembrane #status predicted <TM2>  
F:142-158/Domain: transmembrane #status predicted <TM3>  
F:178-194/Domain: transmembrane #status predicted <TM4>  
F:208-224/Domain: transmembrane #status predicted <TM5>

F:303-319/Domain: transmembrane #status predicted <TM6>  
F:349-365/Domain: transmembrane #status predicted <TM7>  
F:391-407/Domain: transmembrane #status predicted <TM8>  
F:414-430/Domain: transmembrane #status predicted <TM9>

Query Match 47.7%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAMYXXRGI 10  
||| ||:  
Db 129 PAMYMARGL 137

## RESULT 9

probable inner membrane transport protein ECS0007 [imported] - *Escherichia coli* (stra  
C:Species: *Escherichia coli*  
C>Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: G90629  
R:Havashli, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C  
gasaara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic *Escherichia coli* O157:H7 and g  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90629  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <HAY>  
A:Cross-references: GB:BA000007; PIDN:BA83430.1; PID:g13359463; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RIMD 0509952  
C:Genetics:  
A:Gene: ECS0007  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 47.7%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAMYXXRGI 10  
||| ||:  
Db 129 PAMYMARGL 137

## RESULT 10

inner membrane transport protein [imported] - *Escherichia coli* (strain O157:H7, subst  
C:Species: *Escherichia coli*  
C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G85480  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
Iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; DiMantola, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G85480  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <STO>  
A:Cross-references: GB:AE005174; NID:g12512682; PIDN:AG54307.1; GSPDB:GN00145; UMGCP:  
A:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: yaaU  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 47.7%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 PAMYXXRGI 10

Db 129 PAMWARGL 137

## RESULT 11

AC0502  
C:Species: amino-acid transport protein STY0006 [imported] - *Salmonella enterica* subsp. *en*  
C:Species: *Salmonella enterica* subsp. *enterica* serovar Typh  
A:Note: this species has also been called *Salmonella typhi*  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG0502  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,  
th, T.; Cornet, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,  
S.; Moul, S.; O'Gaora, P.  
N:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant *Salmonella enterica* serov  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AG0502  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD01159.1; PID:G16501289; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY0006  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 47.7%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAMWXXRGI 10  
Db 129 PAMWARGL 137

## RESULT 12

A87448  
C:Species: hypothetical protein CC1602 [imported] - *Caulobacter crescentus*  
C:Species: *Caulobacter crescentus*  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: A87448  
R:Merriam, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; J. Lab, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolot  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of *Caulobacter crescentus*.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: A87448  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-545 <STO>  
A:Cross-references: GB:AE005673; NID:913422999; PIDN:AAK2581.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC1602  
C:Superfamily: hypothetical protein u1937b

Query Match 47.7%; Score 41; DB 2; Length 545;  
Best Local Similarity 46.7%; Pred. No. 26;  
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 PAMWXXRGI 16  
Db 242 PCTYRAAGRVPR 256

## RESULT 13

T35841  
C:Species: membrane protein - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999

C:Accession: T35841  
R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, August 1998  
A:Reference number: 221590  
A:Accession: T35841  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-184 <HAR>  
A:Cross-references: EMBL:AL031260; PIDN:CAA20292.1; GSPDB:GN00070; SCORDB:SC9A10.05C  
C:Genetics:  
A:Gene: SCORDB:SC9A10.05C

Query Match 46.5%; Score 40; DB 2; Length 184;  
Best Local Similarity 50.0%; Pred. No. 14;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 3 AMWXXRGI 16  
Db 90 AMWNAHGVQVGF 103

## RESULT 14

T35901  
C:Species: probable arac family transcription regulator - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T35901  
R:Sanders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream,  
submitted to the EMBL Data Library, January 1998  
A:Reference number: 221593  
A:Accession: T35901  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-324 <SAD>  
A:Cross-references: EMBL:AL035212; PIDN:CAA22785.1; GSPDB:GN00070; SCORDB:SC9B2.05  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCORDB:SC9B2.05

Query Match 46.5%; Score 40; DB 2; Length 324;  
Best Local Similarity 57.1%; Pred. No. 24;  
Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 2 PAMWXXRGI 15  
Db 211 PAMWARGL 224

## RESULT 15

G82844  
C:Species: cysteine synthase XE0128 [imported] - *Xylella fastidiosa* (strain 9a5c)  
C:Species: *Xylella fastidiosa*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: G82844  
R:anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq  
N:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: G82844  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-390 <STM>  
A:Cross-references: GB:AE003866; GB:AE003849; NID:9104906; PIDN:AAF82941.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Artuda, P.; Abreu, F.A.; Acencio, M.; Alvarado, R.  
Bionesi, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carier  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr

J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laigh  
 chado, M.A.; Madella, A.M.B.N.; Madella, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, R  
 A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
 , F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
 Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
 A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
 M.; Tsubako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
 A;Reference number: A59328  
 C:Contents: annotation  
 A:Gene: XF0128

Query Match 46.5%; Score 40; DB 2; Length 390;  
 Best Local Similarity 60.0%; Pred. No. 28;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 NPAWYXXRGI 10  
 :|||||  
 :||  
 Db 354 SPAWYAHGI 363

Search completed: September 13, 2002, 09:24:00  
 Job time: 775 sec

THIS PAGE BLANK (USPTO)



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:45 ; Search time 80.21 Seconds

(without alignments)  
8.206 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_5\_21  
Perfect score: 86  
Sequence: 1 NPANVXXRGIRPVGRFX 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	83	96.5	87	1	PRRP_HUMAN
2	82	95.3	83	1	PRRP_RAT
3	82	95.3	98	1	PRRP_BOVIN
4	44	51.2	676	1	EXL1_HUMAN
5	43	50.0	719	1	NRPL_YEAST
6	42	48.8	402	1	EXL1_YEAST
7	41	47.7	476	1	YAAJ_ECOLI
8	39	45.3	149	1	ENRN_BPT7
9	39	45.3	485	1	SYE_BACHD
10	39	45.3	953	1	SYV_VIBCH
11	38.5	44.8	877	1	CAD2_BOVIN
12	38.5	44.8	906	1	CAD2_HUMAN
13	38.5	44.8	906	1	CAD2_MOUSE
14	38.5	44.8	906	1	CAD2_RAT
15	38	44.2	158	1	RL15_AERPE
16	38	44.2	342	1	Y762_METJA
17	38	44.2	347	1	Y576_METJA
18	38	44.2	637	1	MTHA_DROME
19	38	44.2	951	1	SYV_ECOLI
20	38	44.2	1723	1	AIM1_HUMAN
21	38	44.2	2832	1	NDVB_RHIME
22	37	43.0	179	1	REC2_RHIME
23	37	43.0	788	1	REC2_HAEN
24	36	41.9	329	1	Y493_MYCTU
25	36	41.9	334	1	YNHG_ECOLI
26	36	41.9	376	1	OPSL_LIMPO
27	36	41.9	376	1	OPSL_LIMPO
28	36	41.9	994	1	SYV_XYLEA
29	35	40.7	383	1	CYCR_CHRVI
30	35	40.7	385	1	YEIB_ECOLI
31	35	40.7	390	1	RM03_YEAST
32	35	40.7	391	1	THIL_ZOORA
33	35	40.7	442	1	AROA_BORR

34	35	40.7	442	1	AROA_BORR	P12421 bordetella
35	35	40.7	443	1	PR34_CAEEL	O16368 caenorhabd
36	35	40.7	453	1	TBR2_GEOCN	P32925 geophilum
37	35	40.7	456	1	BCS1_YEAST	P32839 saccharomyc
38	35	40.7	460	1	SR54_HAANI	O9hms5 halobacteri
39	35	40.7	596	1	SDP_EIMBO	P42789 elmeria bov
40	35	40.7	618	1	CIR1_CITFR	P23182 citrobacter
41	35	40.7	649	1	RA32_SCHPO	O09683 schizosacch
42	35	40.7	792	1	SYFA_XYLEA	O9pfd6 xylella fas
43	35	40.7	922	1	GYRA_AERSA	P48369 aeromonas s
44	35	40.7	941	1	GCSP_MYCTU	O50601 mycobacteri
45	35	40.7	1035	1	GCSP_SOLTU	O49954 solanum tub

## ALIGNMENTS

```

RESULT 1
PRRP_HUMAN STANDARD; PRT: 87 AA.
ID P81277;
AC P81277;
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 30-MAY-2000 (Rel. 39, last sequence update)
DT 01-MAR-2002 (Rel. 41, last annotation update)
DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing
DE hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-
DE releasing peptide PrRP20].
GN PRH.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RA "A proactin-releasing peptide in the brain.";
RT Nature 393:272-276(1998).
RL
RT
RT
RA Sumino Y., Fujino M.;
RA "Tissue distribution of proactin-releasing peptide (PrRP) and its
RA receptor.";
RT Regul. Pept. 83:1-10(1999).
RL
CC -I- FUNCTION: Stimulates proactin (PRL) release and regulates the
CC expression of proactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -I- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL, AB015419; BAA29027.1; -.
CC DR
CC MIM: 602663; -.
CC KW Hormone; Amladation; Signal.
CC FT
CC FT SIGNAL 1 22
CC FT PEPTIDE 23 53 BY SIMILARITY.
CC FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC FT MOD_RES 53 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC FT SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CMC64;

```

Query Match 96.5%; Score 83; DB 1; Length 87;  
 Best Local Similarity 87.5%; Pred. NO. 1.8e-08;  
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
 ||||| |||||  
 DB 38 NPAMYXXRGIRPVGRF 53

## RESULT 2

PRRP\_RAT STANDARD: PRT: 83 AA.

AC P81278;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP1; Proactin-releasing peptide PrRP2].  
 GN PRI.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OC NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RA "A proactin-releasing peptide in the Brain."  
 RT Nature 393:272-276(1998).  
 RL [2]  
 RP TISSUE SPECIFICITY.

RA PubMed=10498338;  
 RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sunino Y., Fujino M.;  
 RA "Tissue distribution of proactin-releasing peptide (PrRP) and its receptor."  
 RT Regul. Pept. 83:1-10(1999).  
 RL [1]  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR1 EMBL; AB015418; BAA29026.1;  
 KW Hormone; Amidation; Signal;  
 FT SIGNAL 21  
 FT SIGNAL 1  
 FT PEPTIDE 22 52 PROACTIN-RELEASING PEPTIDE PRRP1.  
 FT PEPTIDE 33 52 PROACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 32 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).  
 SO SEQUENCE 83 AA; 9215 MW; D0C75A264EBE4F29 CRC64;

Query Match 95.3%; Score 82; DB 1; Length 83;  
 Best Local Similarity 87.5%; Pred. NO. 2.7e-08;  
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
 ||||| |||||  
 DB 37 NPAMYXXRGIRPVGRF 52

RESULT 3  
 PRRP\_BOVIN STANDARD: PRT: 98 AA.

AC P81264;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP1; Proactin-releasing peptide PrRP2].  
 GN PRI.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RC TISSUE-Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RA "A proactin-releasing peptide in the Brain."  
 RT Nature 393:272-276(1998).  
 RL [1]  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC DR EMBL; AB015417; BAA29025.1;  
 KW Hormone; Amidation; Signal;  
 FT SIGNAL 22  
 FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP1.  
 FT PEPTIDE 33 53 PROACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SO SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 95.3%; Score 82; DB 1; Length 98;  
 Best Local Similarity 87.5%; Pred. NO. 3.1e-08;  
 Matches 14; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 NPAMYXXRGIRPVGRF 16  
 ||||| |||||  
 DB 38 NPAMYXXRGIRPVGRF 53

## RESULT 4

EX1L\_HUMAN STANDARD: PRT: 676 AA.

AC Q92935;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Exostosin-like 1 (Exostosin-L1) (Multiple exostosin-like protein).  
 GN EX1L OR EXT1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RX MEDLINE=97189339; PubMed=9037597;

```

RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
RT "Identification and localization of the gene for EXT_L, a third member  
RL of the multiple exostosin gene family.";  
RN Genome Res. 7;10-16(1997).  
[2]  
RP SEQUENCE FROM N.A.  
RQ Xu L., Deng H.X., Xia J.H., Pan O., Liu C.Y.;  
RA "Mutations of the EXT genes in hereditary multiple exostoses in  
RT Chinese";  
RL Submitted (AUG-1998) to the EMBL/GenBank/DDBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RP Mytts W., Spleker N., Van Roy N., De Paeppe A., De Bouille K.,  
RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;  
RT "Refined physical mapping and genomic structure of the EXT_L1 gene.";  
RL Submitted (MAY-1999) to the EMBL/Genbank/DDJB databases.  
CC -I FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
CC -I SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
CC reticulum (By similarity).  
CC -----  
CC -I SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See http://www.isdb.ch/announce/  
CC or send an email to licens@isdb-slb.ch).  
-----  
DR EMBL: U67191; AACSI141.1; "-"  
DR EMBL: AF083653; AAD02840.1; "-"  
DR EMBL: AF083653; AAD02840.1; JOINED.  
DR EMBL: AF083624; AAD02840.1; JOINED.  
DR EMBL: AF083625; AAD02840.1; JOINED.  
DR EMBL: AF083626; AAD02840.1; JOINED.  
DR EMBL: AF083627; AAD02840.1; JOINED.  
DR EMBL: AF083628; AAD02840.1; JOINED.  
DR EMBL: AF083629; AAD02840.1; JOINED.  
DR EMBL: AF083630; AAD02840.1; JOINED.  
DR EMBL: AF083631; AAD02840.1; JOINED.  
DR EMBL: AF083632; AAD02840.1; JOINED.  
DR EMBL: AF153960; AAFF3172.1; "-"  
DR EMBL: AF153961; AAFF3172.1; JOINED.  
DR MIR: 601738; "-"  
DR InterPro: IPR004263; Exostosin.  
DR Pfam: PF03016; Exostosin; 1.  
KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.  
FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
         (POTENTIAL).  
FT CARBOHYD 269 269 N-LINKED (GLCNAAC... ) (POTENTIAL).  
SQ SEQUENCE 676 AA; 74673 MM; BSE006AB76ZDE5633 CRC64;
```

CC	Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OC	NCBI_TaxID=4932;
RM	SEQUENCE FROM N.A.
RP	STRAIN-AH22;
RC	MEDLINE=93247548; PubMed=8483449;
RA	Wehner E.P., Rao E., Brendel M.;
RT	"Molecular structure and genetic regulation of SFA, a gene responsible for resistance to formaldehyde in <i>Saccharomyces cerevisiae</i> , and characterization of its protein product.";
RL	Mo. Gen. Genet. 237:351-358(1993).
RN	[2]
RP	SEQUENCE FROM N.A.
RC	STRAIN-S288C;
RA	Pohl T.M.;
RL	Submitted (NCV-1995) to the EMBL/GenBank/DBJ databases.
CC	-1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.
CC	-1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).
CC	-----
DR	EMBL; X68020; CAA48159.1; -;
DR	EMBL; Z67750; CAA91579.1; -;
DR	EMBL; Z74215; CAA98741.1; -;
DR	PIR; S31139; S31139.
DR	HSSP; P04170; GRXN.
DR	SGD; S0002326; NRPI.
DR	InterPro; IPR00504; RRM.
DR	InterPro; IPR001876; ZnF-RanBP.
DR	Pfam; PF00076; rtm: 1
DR	Pfam; PF00641; Zf-RanBP; 2.
DR	SMART; SM00360; RRM; 1.
DR	SMART; SM00547; ZnF_RBZ; 2.
DR	PROSITE; PS50102; RRM; 1.
DR	PROSITE; PS00030; RRM_RNP_1; FALSE_NEG.
DR	PROSITE; PS01358; ZF_RANBP_1; 2.
DR	PROSITE; PS50193; ZF_RANBP2_2; 2.
RW	Nuclear protein; Zinc-finger; RNA-binding; Repeat.
FT	DOMAIN 226 322
FT	ZN_FING 355 384
FT	ZN_FING 581 610
FT	DOMAIN 490 564
FT	DOMAIN 493 493
FT	CONFLICT 493 493 I -> N (IN REF. 1).
SO	SEQUENCE 719 AA; 79299 MW; ADA9BC09FD582669 CRC64;
-----	
Query Match 50.0%; Score 43; DB 1; Length 719;	
Best Local Similarity 50.0%; Pred. No. 3.4;	
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;	
QY	3 AWYXXRGIRPYG 14
DB	1: 1:
DB	244 SMFTGYGVRPYG 255
-----	
RESULT 6	
EX7L_STRCO	
ID EX7L_STRCO	STANDARD; PRT; 402 AA.
AC	Q9FBM3;
DT	16-OCT-2001 (Rel. 40, Created)
DT	16-OCT-2001 (Rel. 40, Last sequence update)
DT	16-OCT-2001 (Rel. 40, Last annotation update)
DE	Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)
DE	(Exonuclease VII large subunit).
GN	XSEFA OR SCK7.29C.
OS	Streptomyces coelicolor.
OS	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;

```

AC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomycetes.
OX NCBI_TaxId:1902;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NA3(2);
RA Seeger K.J., Harris D., Cerdano A.M., Parkhill J., Barrell B.G.,
RA Rajandream M.A.;
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.
CC -!- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).
CC -!- CATALYTIC ACTIVITY: Exonuclease activity: cleavage in either 5'- to 3'-
CC or 3'- to 5'-direction to yield 5'-phosphonucleotides.
CC -!- SUBUNIT: HETERODIMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY
CC SIMILARITY).
CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -!- SIMILARITY: BELONGS TO THE XSEA FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AL391754; CAC05901.1; -
DR InterPro; IPR003753; Exonuc_VII_L.
DR InterPro; IPR002309; tRNA-synt_2.
DR Pfam; PF02601; Exonuc_VII_L; 1.
DR Pfam; PF01336; tRNA_antl; 1.
KW Hydrolyase; Nuclease; Exonuclease.
SO SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 48.8%; Score 42; DB 1; Length 402;
Best Local Similarity 42.9%; Pred. No. 2.9;
Matches 9; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

OY 2 PAMVXXRG-----IRPVG 14
   1 11 11 1:111
Db 89 PEMVAPRGQLSLRAAEIKRPVG 109

RESULT 7
YAAJ_ECOLI STANDARD; PRT; 476 AA.
AC P30143;
ID YAAJ_ECOLI STANDARD; PRT; 476 AA.
DT 01-JUL-1993 (Rel. 26, Created)
DT 01-JUL-1993 (Rel. 26, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Putative transporter yaaJ.
GN YAAJ OR B0007.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxId:562;
OX [1]
RN RN
RP RP
RC RC
RA RA
RA Yura T., Morl H., Nagai H., Nagata T., Ishihama A., Fujita N.,
RA Isono K., Mizobuchi K., Nakata A.;
RT "Systematic sequencing of the Escherichia coli genome: analysis of
RT the 0-2.4 min region."
RL Nucleic Acids Res. 20:3305-3308(1992).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RC MEDLINE=97426617; PubMed=9278503;
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

Query	2	PAMXXXXGCI 10	47.7%; Score 41; DB 1; Length 476;
Best Local Similarity	66.7%;	Pred. No. 5.3;	
Matches	6; Conservative	1; Mismatches	2; Indels
			0; Gaps
			0;
Db	129	PAMWARGL 137	
QY	2	PAMXXXXGCI 10	
RESULT	8		
ENRN_BPT7			
ID	ENRN_BPT7	STANDARD;	PRT; 149 AA.
AC	P00641;		
DT	21-JUL-1986 (Rel. 01, Created)		
DT	21-JUL-1986 (Rel. 01, Last sequence update)		
DT	01-NOV-1995 (Rel. 32, Last annotation update)		
DE	Endonuclease I (EC 3.1.21.2) (Endonuclease).		
GN	3.		
OS	Bacteriophage T7.		
OC	Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;		
OC	T7-like phages.		
OX	NCBI_TaxID=10760;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=83241725; PubMed=6864790;		
RA	Dunn J.J., Studier F.W.;		
RT	"Complete nucleotide sequence of bacteriophage T7 DNA and the		
RT	locations of T7 genetic elements.";		
RL	J. Mol. Biol. 166:477-535(1983).		
RN	[2]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=82078034; PubMed=7310871;		
RA	Dunn J.J., Studier F.W.;		
RT	"Nucleotide sequence from the genetic left end of bacteriophage T7		

```

RT DNA to the beginning of gene 4."
RL J. Mol. Biol. 148:303-330(1981).
CC -1- FUNCTION: ENODOXYRIBONUCLEASE I, WHICH IS EXPRESSED IN THE LATE
CC STAGE, IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN
CC OF HOST DNA. IN THE EARLY STAGE OF INFECTION, T7 DNA REPLICATES AS
CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA
CC LINEAR CONCATEMERS SEVERAL GENOMES IN LENGTH. THE GENE 3 PRODUCT
CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE GENOMES.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'.
CC Phosphooligonucleotide end-products.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: V01127; CAA24345.1; -
DR EMBL: V01146; CAA24402.1; -
DR PIR: A00785; NEBP37.
DR PIR: S42301; S42301.
KM Hydrolyase: Nuclease: Endonuclease.
SQ SEQUENCE 149 AA; 17172 MW; D092MA28E3743BC1 CRC64;

Query Match
Best Local Similarity 45.3%; Score 39; DB 1; Length 149;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 3 AMYXRGIRPVGRF 16
DB 2 AGYGAKGIRKVGAF 15

RESULT 9
STE_BACHD STANDARD; PRT; 485 AA.
ID SYE_BACHD
AC 09KGF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamy1-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
DE (GluRS).
GN GUTX OR BH0109.
OS Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86655;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fujii F., Hirano C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.",
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) -> AMP +
CC diphosphate + L-glutamyl-tRNA(Glu).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: AP001507; BAB03828.1; -
DR InterPro: IPR000924; tRNA-synt_1c.
DR InterPro: IPR001412; tRNA-synt_1.
DR Pfam: PR00749; tRNA-synt_1c; 1.
DR PRINTS: PR00987; TRANSYNTHLU.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 11 21 "HIGH" REGION.
FT SITE 252 256 "KMSKS" REGION.
FT BINDING 255 255 ATP (BY SIMILARITY).
SQ SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match
Best Local Similarity 45.3%; Score 39; DB 1; Length 485;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 5 YXXRGIRPVGRF 16
DB 146 YEAKGIRKPVGRF 157

RESULT 10
STE_VIBCH STANDARD; PRT; 953 AA.
ID SYV_VIBCH
AC 09KPF3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Valyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (ValRS).
GN VALS OR VC2503.
OS Vibrrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gilm M.L.,
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Yamahirov J., Bass S., Qin H., Dreigol I., Sellers P.,
RA McDonald L., Utterback T., Fleischmann R.D., Mierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) -> AMP + diphosphate
CC + L-valyl-tRNA(Val).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AF004320; AAF95645.1; -
DR HSP: P96142; IGAX.
DR TIGR: VC2503; -
DR InterPro: IPR002300; tRNA-synt_1a.
DR InterPro: IPR001412; tRNA-synt_1.
DR InterPro: IPR002303; tRNA-synt_val.
DR Pfam: PR00133; tRNA-synt_1; 1.
DR PRINTS: PR00986; TRANSYNTHVAL.
DR PROSITE: PS00178; AA_TRNA_LIGASE_I; 1.

```

KW	Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
RN	Complete proteome.
FT	SITE 42 "HIGH" REGION.
FT	SITE 554 "KMSKS" REGION.
FT	BINDING 557 ATP (BY SIMILARITY).
SO	SEQUENCE 953 AA; 108170 MW; D93471A33CF4F69C CRC64;
OY	2 PAMVXXRGIRPYGR 15
DB	436 PAMVDEQGNVEVGR 449
Query Match	45.3%; Score 39; DB 1; Length 953;
Best Local Similarity	57.1%; Pred. No. 25;
Matches	8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;
RESULT 11	
CAD2_BOVIN	STANDARD; PRT; 877 AA.
ID	P19534;
AC	01-FEB-1991 (Rel. 17, created)
DT	01-FEB-1991 (Rel. 17, last sequence update)
DT	16-OCT-2001 (Rel. 40, last annotation update)
DE	Neural-cadherin precursor (N-cadherin) (Cadherin-2) (Fragment). CDH2.
OS	Bos taurus (Bovine).
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
CC	Bovidae; Bovinae; Bos. NCBI_TaxID=9913;
RN	[1]
RP	SEQUENCE FROM N.A.
RX	MEDLINE=90360979; PubMed=2390969;
RA	Lisaw C.W., Cannon C., Power M.D., Kiboneka P.K., Rubin L.L.?
RT	"Identification and cloning of two species of cadherins in bovine endothelial cells."
RL	EMBO J. 9:2701-2708(1990).
CC	-1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS. THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN NEURONAL RECOGNITION MECHANISMS.
CC	-1- SUBCELLULAR LOCATION: Type I membrane protein.
CC	-1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see http://www.isb-sdb.ch/announce/ or send an email to license@sdb.ch).
CC	-----
DR	EMBL; X53615; CAA37677.1; .
DR	PIR; S11693; IJBOCN.
DR	HSSP; P15116; INCI.
DR	InterPro; IPRO002126; Cadherin.
DR	InterPro; IPRO00233; Cadherin_C_term.
DR	Planm; PF00028; cadherin_5.
DR	Planm; PF01049; cadherin_C_term; 1.
DR	PRINTS; PR00205; CADHERIN.
DR	SMART; SM00112; CA; 5.
DR	PROSITE; PS00232; CADHERIN_1; 3.
DR	PROSITE; PS00268; CADHERIN_2; 5.
KW	Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.
FT	NON_TER 1
FT	PROPEP 1
FT	<1 130
FT	CHAIN 131 877 NEURAL-CADHERIN.
FT	DOMAIN 131 695 EXTRACELLULAR (POTENTIAL).
FT	TRANSMEM 696 717 POTENTIAL.
FT	CYTODASMIC 877 CYTOPLASMIC (POTENTIAL).
FT	DOMAIN 131 238 CADHERIN 1.

[illegible]

CC NEURONAL RECOGNITION MECHANISM.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: X57548; CAA40773.1; -  
 DR EMBL: X54315; CAA38213.1; -  
 DR EMBL: S42303; AAB22854.1; -  
 DR EMBL: M34064; AAA03236.1; -  
 DR EMBL: 227420; CAA81799.1; -  
 DR PIR: A38870; IJHUCN.  
 DR HSSP: P15116; INCU.  
 DR MIM: 114020; -  
 DR InterPro: IPR002126; Cadherin.  
 DR InterPro: IPR000233; Cadherin\_C-term.  
 DR Pfam: PF00028; cadherin\_5.  
 DR Pfam: PF01049; cadherin\_C-term; 1.  
 DR PRINTS: PR00205; CADHERIN.  
 DR PROSITE: PS00232; CADHERIN\_1; 3.  
 DR PROSITE: PS50268; CADHERIN\_2; 5.  
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KM Signal.  
 FT SIGNAL. 1 23  
 FT PROPEP 24 159  
 FT CHAIN 160 906  
 FT DOMAIN 160 724  
 FT TRANSMEM 725 746  
 FT DOMAIN 747 906  
 FT DOMAIN 160 267  
 FT DOMAIN 267 382  
 FT DOMAIN 383 497  
 FT DOMAIN 498 603  
 FT DOMAIN 604 714  
 FT DOMAIN 714 878  
 FT CARBOHYD 190 273  
 FT CARBOHYD 325 402  
 FT CARBOHYD 402 462  
 FT CARBOHYD 572 651  
 FT CARBOHYD 651 692  
 FT CARBOHYD 692 712  
 FT CONFLICT 16 16  
 FT CONFLICT 196 196  
 FT CONFLICT 212 212  
 FT CONFLICT 357 357  
 FT CONFLICT 867 867  
 FT SEQUENCE 906 AA; 99851 MW; 72DDC7B8B57C7AFC CRC64;  
 Query Match 44.8%; Score 38.5; DB 1; Length 906;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Neuronal cadherin precursor (N-cadherin) (Cadherin-2).  
 GN CDH2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RX MEDLINE=99346748; PubMed=2762814;  
 RA Miyatani S., Shimamura K., Hatta M., Nagafuchi A., Nose A.,  
 RA Matsunaga M., Hatta K., Takeichi M.;  
 RT "Neural cadherin: role in selective cell-cell adhesion.";  
 RL Science 245:631-635(1989).  
 RP SEQUENCE FROM N.A.  
 RA Tamura K.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN PARTIAL SEQUENCE FROM N.A.  
 RP STRAIN=C57BL/6;  
 RC MEDLINE=92409532; PubMed=1528849;  
 RA Miyatani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;  
 RT "Genomic structure and chromosomal mapping of the mouse N-cadherin  
 RT gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8443-8447(1992).  
 RN [4]  
 RP DEVELOPMENTAL STAGE.  
 RC STRAIN=C57BL/6; TISSUE=Testis;  
 RX MEDLINE=97033837; PubMed=8879495;  
 RA Munro S.B., Blaschuk O.W.;  
 RT "A comprehensive survey of the cadherins expressed in the testes of  
 RT fetal, immature, and adult mice utilizing the polymerase chain  
 RT reaction.";  
 RL Biol. Reprod. 55:822-827(1996).  
 RN [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 160-267.  
 RX MEDLINE=95191680; PubMed=7885471;  
 RA Shapiro L., Fannon A.M., Kwong P.D., Thompson A., Lehmann M.S.,  
 RA Grubel G., Legrand J.-F., Als-Nielsen J., Colman D.R.,  
 RA Hendrickson W.A.;  
 RT "Structural basis of cell-cell adhesion by cadherins.";  
 RL Nature 374:327-337(1995).  
 RN [6]  
 RP X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS) OF 160-374.  
 RX MEDLINE=98318235; PubMed=9655503;  
 RA Tamura K., Shan W.S., Hendrickson W.A., Colman D.R., Shapiro L.;  
 RT "Structure-function analysis of cell adhesion by neural (N-)  
 RT cadherin.";  
 RL Neuron 20:1153-1163(1998).  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
 CC NEURONAL RECOGNITION MECHANISM.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF TESTICULAR  
 CC DEVELOPMENT WITH HIGHEST LEVELS FOUND IN TESTES OF 21-DAY-OLD  
 CC MICE.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC at the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: M31131; AAA7353.1; -  
 DR EMBL: AB008811; BAA23549.1; -  
 DR EMBL: S45011; AAB23356.1; -  
 DR PIR: A32759; IJMSCN.

DR PDB; INCG; 10-JUL-95.  
 DR PDB; INCH; 10-JUL-95.  
 DR PDB; INCI; 10-JUL-95.  
 DR PDB; INCI; 18-MAR-99.  
 DR MGI; 88355; Cdh2.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR002233; Cadherin\_C-term.  
 DR Pfam; PF00028; cadherin; 5.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS0268; CADHERIN\_2; 5.  
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal; 3D-structure.  
 FT SIGNAL 1 23  
 FT PROPEP 24 159  
 FT CHAIN 160 906  
 FT DOMAIN 160 724  
 FT TRANSMEM 725 746  
 FT DOMAIN 747 906  
 FT DOMAIN 160 267  
 FT DOMAIN 268 382  
 FT DOMAIN 383 497  
 FT DOMAIN 498 603  
 FT DOMAIN 604 717  
 FT DOMAIN 717 863  
 FT CARBOHYD 190 190  
 FT CARBOHYD 273 273  
 FT CARBOHYD 325 325  
 FT CARBOHYD 402 402  
 FT CARBOHYD 572 572  
 FT CARBOHYD 651 651  
 FT CARBOHYD 692 692  
 FT CONFLICT 7 9  
 FT CONFLICT 565 565  
 FT CONFLICT 567 567  
 FT CONFLICT 624 624  
 FT SEQUENCE 906 AA; 99761 MW; 0206741B77107B7 CRC64;  
 Query Match 44.8%; Score 38.5; DB 1; Length 906;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
 QY 2 PAM--XXRGIRPYGRF 16  
 DB 414 PAMNAAYRISGDPYGRF 431  
 RESULT 14  
 CAD2\_RAT STANDARD; PRT; 906 AA.  
 AC 0981Y3: O9R0R5;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE Neutral-cadherin precursor (N-cadherin) (Cadherin-2).  
 GN Cdh2.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
 RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Testis;  
 RX MEDLINE=98187820; PubMed=9528971.  
 RA Chung S.S., Mo M.Y., Silvestri B., Lee W.M., Cheng C.Y.;  
 RT "Rat testicular N-cadherin: its complementary deoxyribonucleic acid  
 cloning and regulation.";  
 RL Endocrinology 139:1853-1862(1998).  
 RP SEQUENCE FROM N.A.

RC STRAIN-SPRAGUE DAWLEY; TISSUE-Brain;  
 RA Asai K., Tada T., Yamamoto M., Tada A., Mizuno M., Eimoto T., Kato T.;  
 RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
 CC NEURONAL RECOGNITION MECHANISM.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
 CC -1- TISSUE SPECIFICITY: IN TESTIS, EXPRESSED IN SEROTOLI AND GERM  
 CC CELLS.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; AF097593; AAC83818.1; -;  
 DR EMBL; AB017695; BAAB4919.1; -;  
 DR HSSP; P15116; INCI.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR00233; Cadherin\_C-term.  
 DR Pfam; PF00028; cadherin; 5.  
 DR Pfam; PF01049; Cadherin\_C-term; 1.  
 DR PRINTS; PR00205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS0268; CADHERIN\_2; 5.  
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KW Signal.  
 FT SIGNAL 1 27  
 FT PROPEP 28 159  
 FT CHAIN 160 906  
 FT DOMAIN 160 724  
 FT TRANSMEM 725 745  
 FT DOMAIN 746 906  
 FT DOMAIN 160 267  
 FT DOMAIN 268 382  
 FT DOMAIN 383 497  
 FT DOMAIN 498 603  
 FT DOMAIN 604 717  
 FT DOMAIN 717 863  
 FT CARBOHYD 190 190  
 FT CARBOHYD 273 273  
 FT CARBOHYD 325 325  
 FT CARBOHYD 402 402  
 FT CARBOHYD 572 572  
 FT CARBOHYD 622 622  
 FT CARBOHYD 651 651  
 FT CARBOHYD 692 692  
 FT CONFLICT 7 9  
 FT CONFLICT 48 48  
 FT CONFLICT 153 153  
 FT CONFLICT 646 646  
 FT CONFLICT 658 658  
 FT CONFLICT 724 724  
 FT SEQUENCE 906 AA; 99685 MW; 97D9937FD8D1F8B5 CRC64;  
 Query Match 44.8%; Score 38.5; DB 1; Length 906;  
 Best Local Similarity 50.0%; Pred. No. 29;  
 Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;  
 QY 2 PAM--XXRGIRPYGRF 16  
 DB 414 PAMNAAYRISGDPYGRF 431  
 RESULT 15



```

RL15_AERPE
ID RL15_AERPE STANDARD; PRT; 158 AA.
AC Q9YF98;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 50S ribosomal protein L15P.
GN RPL15P OR APE0343.
OS Aeropyrum pernix.
OC Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
OC Aeropyrum.
OX NCBI_TaxID=56636;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K1;
RX MEDLINE=99310339; PubMed=10382966;
RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA Jin-no K., Takahashi M., Sekine M., Baba S.-I., Anai A., Kosugi H.,
RA Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT *Complete genome sequence of an aerobic hyper-thermophilic
RT crenarchaeon, Aeropyrum pernix K1."
RL DNA Res. 6:83-101(1999).
CC -!- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AP000059; BAA79298.1;
DR InterPro: IPR001196; Ribosomal_L15.
DR PROSITE: PS00475; RIBOSOMAL_L15; 1.
KW Ribosomal protein; Complete proteome.
SQ SEQUENCE 158 AA; 17480 MW; 01220322DA157321 CRC64;

Query Match 44.2%; Score 38; DB 1; Length 158;
Best Local Similarity 45.5%; Pred. No. 6.2;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PAVYXXRGIRP 12
DB 54 PTWYGRKGFNP 64

```

Search completed: September 13, 2002, 09:30:46  
Job time: 1136 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:20 ; Search time 311.85 Seconds  
(without alignments)  
9.431 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_5\_21  
Perfect score: 86  
Sequence: 1 NPAWYXXRGIRPVGRFX 17

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues  
Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	68	79.1	117	13	Q9W624
2	46.5	54.1	420	10	Q93L27
3	46	53.5	790	10	Q9M371
4	44	51.2	333	16	Q9PH76
5	44	51.2	419	4	Q9Y276
6	43	50.0	220	16	Q91022
7	43	50.0	418	11	Q9CZP5
8	42	48.8	433	16	Q9A382
9	41	47.7	338	5	Q18729
10	41	47.7	545	16	Q9A7W7
11	40	46.5	153	9	Q9T133
12	40	46.5	184	2	Q86838
13	40	46.5	324	2	Q9Z554
14	40	46.5	390	16	Q9PH18
15	40	46.5	486	16	Q9ZDM2
16	40	46.5	540	10	Q9LGZ0

17	40	46.5	562	5	Q9VN94	Q9VN94 drosophila
18	39	45.3	97	2	Q33440	Q33440 pseudomonas
19	39	45.3	108	11	Q9D1V4	Q9D1V4 mus musculus
20	39	45.3	179	2	Q9K307	Q9K307 streptomyces
21	39	45.3	256	16	Q9ZVE9	Q9ZVE9 rhizobium m
22	39	45.3	327	10	Q9LU14	Q9LU14 arabidopsis
23	39	45.3	333	4	Q96SD4	Q96SD4 homo sapien
24	39	45.3	340	5	Q18731	Q18731 caenorhabdi
25	39	45.3	370	11	Q9D3V7	Q9D3V7 mus musculu
26	39	45.3	416	16	Q99Z49	Q99Z49 streptococc
27	39	45.3	425	16	Q98606	Q98606 rhizobium l
28	39	45.3	555	3	Q00050	Q00050 aspergillus
29	39	45.3	538	10	Q9ZVE3	Q9ZVE3 arabidopsis
30	39	45.3	2104	5	Q21281	Q21281 caenorhabdi
31	39	45.3	2104	5	Q964N4	Q964N4 caenorhabdi
32	38.5	44.8	328	11	Q55075	Q55075 cricetus
33	38.5	44.8	360	5	Q19879	Q19879 caenorhabdi
34	38	44.2	127	16	Q9HXA7	Q9HXA7 pseudomonas
35	38	44.2	178	16	Q97SU7	Q97SU7 streptococc
36	38	44.2	194	16	Q9A9U0	Q9A9U0 caulobacter
37	38	44.2	210	10	Q9LEM6	Q9LEM6 catharanthu
38	38	44.2	226	16	Q9A359	Q9A359 caulobacter
39	38	44.2	238	17	Q9YEH0	Q9YEH0 aeropyrum p
40	38	44.2	359	2	Q88036	Q88036 streptomyces
41	38	44.2	398	16	Q9RR92	Q9RR92 delnoccocus
42	38	44.2	481	5	Q95WT7	Q95WT7 drosophila
43	38	44.2	536	5	Q95YN4	Q95YN4 drosophila
44	38	44.2	536	5	Q95PK7	Q95PK7 drosophila
45	38	44.2	536	5	Q95PK6	Q95PK6 drosophila

## ALIGNMENTS

RESULT 1	
ID Q9W624	PRELIMINARY; PRT; 117 AA.
AC Q9W624;	
DT 01-NOV-1999 (TREMBlrel. 12, Created)	
DT 01-NOV-1999 (TREMBlrel. 12, Last sequence update)	
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)	
DE C-RF AMIDE.	
OS Carassius auratus (Goldfish).	
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;	
OC Cypriniformes; Cyprinidae; Carassius.	
OX NCBI_TaxId=7957;	
RN (1)	
RP SEQUENCE FROM N.A.	
RC TISSUE=BRAIN;	
RA Satake H., Minakata H., Fujimoto M.;	
RT "Carassius Rhamde (C-RF amide).";	
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.	
DR EMBL; AB020024; BAA76662.1; -	
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;	

Query Match	79.14;	Score 68;	DB 13;	Length 117;
Best Local Similarity	62.5%;	Pred. NO. 0.00011;		
Matches 10;	Conservative	3;	Mismatches 3;	Indels 0;
QY	1 NPAWYXXRGIRPVGRF 16			
	:			
Db	60 DPMYVGRGVPIGRF 75			
RESULT	2			
Q93L27				
ID	Q93L27	PRELIMINARY;	PRT;	420 AA.
AC	Q93L27;			
DT	01-DEC-2001 (TREMBlrel. 19, Created)			
DT	01-DEC-2001 (TREMBlrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBlrel. 19, Last annotation update)			

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN ADR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
CX NCBI\_TaxID=1694;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CCM3239;  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CCM3239."  
RL Submitted (MAY-2001) to the EMBL/Genbank/DBJ databases.  
DE EMBL: AY033994; AAK61719.1; -  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E2BE8C2DEA CRC64;  
  
Query Match 54.1%; Score 46.5; DB 2; Length 420;  
Best Local Similarity 52.6%; Pred. No. 3.4;  
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;  
  
OY 3 AMYXX-----RGIRPVGRF 16  
||: |||||  
DB 37 AMMAVNLGSGIRPVGRF 55  
  
RESULT 3  
ID 09M371 PRELIMINARY; PRT: 790 AA.  
AC 09M371.  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYPOTHEICAL 87.4 KDA PROTEIN.  
GN F15G16.60.  
OS Arabidopsis thaliana (mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
CC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eustosids II; Brassicales; Brassicaceae; Arabidopsids.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De Haan M., Maatse A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Quetier F., Salanoubat M.;  
RL Submitted (NOV-1999) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (JAN-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL: AL132959; CAB71097.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;  
  
Query Match 53.5%; Score 46; DB 10; Length 790;  
Best Local Similarity 50.0%; Pred. No. 8.2;  
Matches 8; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
  
OY 1 NPAWXXRGIRPVGRF 16  
||: |||||  
DB 369 NPTYGSRGLDPHGRW 384  
  
RESULT 4  
ID 09PH76 PRELIMINARY; PRT: 333 AA.  
AC 09PH76.  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE HYDROXYBENZONATE OCTAPRENYLTRANSFERASE.  
GN XF0068.  
OS Xylella fastidiosa.

OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
CX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Bata G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Biondes M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
RA Coutinho N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohelsel J.D., Junqueira M.L., Kemper E.L., Kitaajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madelira A.M.B.N., Madelira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000)  
DR EMBL: AE003860; AAE02881.1; -  
DR InterPro: IPR000537; UDLA.  
DR Pfam: PF01040; UDLA. 1.  
DR PROSITE: PS00943; UDLA; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 333 AA; 37931 MW; ECF3F4716C962B95 CRC64;  
  
Query Match 51.2%; Score 44; DB 16; Length 333;  
Best Local Similarity 57.1%; Pred. No. 7.3;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
  
OY 1 NPAWXXRGIRPVG 14  
||: |||||  
DB 55 DPYKRLAGDRPVG 68  
  
RESULT 5  
ID 09Y276 PRELIMINARY; PRT: 419 AA.  
AC 09Y276.  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE H-BCSI (BCSI (YEAST HOMOLOG)-LIKE).  
GN BCSI OR BCSIL.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
CX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RX MEDLINE=99097350; PubMed=9878253;  
RA Petruzzella V., Tiranti V., Fernandez P., Ianna P., Carozzo R.,  
RA Zeviani M.;  
RT "Identification and characterization of human cDNAs specific to BCSI,  
RT PET112, SCOL, COX15, and COXII, five genes involved in the formation

RT and function of the mitochondrial respiratory chain.";  
RL Genomics 54:494-504(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE-96207227; PubMed-8619474;  
RA Anderson B., Wentland M.A., Ricafrente J.Y., Liu W., Gibbs R.A.;  
RT "A double adaptor" method for improved shotgun library  
RT construction.";  
RL Anal. Biochem. 236:107-113(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RX MEDLINE-97264341; PubMed-9110174;  
RA Yu W., Anderson B., Morley K.C., Muzny D.M., Ding Y., Liu W.,  
RA Ricafrente J.Y., Wentland M.A., Lennon G., Gibbs R.A.;  
RT "Large-scale concatenation cDNA sequencing.";  
RL Genome Res. 7:353-358(1997).  
RN [4]  
RP SEQUENCE FROM N.A.  
RA de Lontay P., Valnot I., Barrientos A., Gorbalyuk M., Tzagoloff A.,  
RA Benayon E., Chretien D., Kadom N., Lombes A., Ogier de Baulny H.,  
RA Nludet P., Munnich A., Rustin P., Rotig A.;  
RT "Mutations in bcl, a mitochondrial respiratory chain assembly gene,  
RT are responsible for the complex III deficiency of patients with  
RT tubulopathy and liver failure.";  
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
RN [5]  
RP SEQUENCE FROM N.A.  
RC TISSUE-RHABDOMYOSARCOMA;  
RA Strausberg R.;  
RL Submitted (NOV-2000) to the EMBL/GenBank/DBJ databases.  
RN [6]  
RP SEQUENCE FROM N.A.  
RC TISSUE-MUSCLE, RHABDOMYOSARCOMA;  
RA Strausberg R.;  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF026849; AAD08638.1; -  
DR EMBL; AF038195; AAB97365.1; -  
DR EMBL; AF346835; AAK29417.1; -  
DR EMBL; BC000416; AAH00416.1; -  
DR EMBL; BC07500; AAH07500.1; -  
DR InterPro: IPR003593; AAA;  
DR InterPro: IPR003959; AAA\_subfam.  
DR Pfam: PF00004; AAA; 1.  
DR SMART: SM00382; AAA; 1.  
KM ATP-binding; Hypothetical protein.  
SQ SEQUENCE 419 AA; 47534 MW; 7F0F98BA62F2CBB8 CRC64;

Query Match 51.2%; Score 44; DB 4; Length 419;  
Best Local Similarity 70.0%; Pred. No. 9.4;  
Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 NPAAVXXRCI 10  
|||  
DB 211 NPKMYTDRGI 220

RESULT 6  
ID 091022 PRELIMINARY; PRT; 220 AA.  
AC 091022;  
DT 01-MAR-2001 (TREMblrel. 16, Created)  
DT 01-MAR-2001 (TREMblrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE PROBABLE GLUTATHIONE S-TRANSFERASE.  
GN PA2821.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
NCBI\_TaxID=287;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 15692 / PA01;  
RX MEDLINE-20437337; PubMed-10984043;  
RA Stover C.K., Pham X.-O.T., Ervin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,  
RA Gardner R.L., Goltzy L., Tolentino E., Westbrock-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL; AE004709; AAG06209.1; -  
DR InterPro: IPR004046; GST\_C.  
DR InterPro: IPR004045; GST\_N.  
KM Transferase; Complete proteome.  
SQ SEQUENCE 220 AA; 24716 MW; 6596183EACAA050 CRC64;

Query Match 50.0%; Score 43; DB 16; Length 220;  
Best Local Similarity 64.3%; Pred. No. 7.1;  
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

OY 2 PAAVXXRGIRPVR 15  
|||||  
DB 39 PAAV--REISPLGR 50

RESULT 7  
ID 09CZP5 PRELIMINARY; PRT; 418 AA.  
AC 09CZP5;  
DT 01-JUN-2001 (TREMblrel. 17, Created)  
DT 01-JUN-2001 (TREMblrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE 9130022019RIK PROTEIN.  
GN 9130022019RIK  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C57BL/6J; TISSUE-EMBRYO;  
RX MEDLINE-21085660; PubMed-11217651;  
RA Kawai J., Shingawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arakawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamana K.I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasukawa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Glasl C., King B., Kochia H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schiraldi L.M., Staudt F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsch G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hume D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyokura K., Wang K.H., Weitz C., Whitaker C., Wilming L.,  
RA Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kohlsuk S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection.";  
RL Nature 409:685-690(2001).  
DR EMBL; AK012324; BAB28162.1; -  
DR MGD; MGI:1914071; 9130022019RIK.  
DR InterPro: IPR003593; AAA.  
DR InterPro: IPR003959; AAA\_subfam.  
DR Pfam: PF00004; AAA; 1.  
DR SMART: SM00382; AAA; 1.  
KM ATP-binding.  
SQ SEQUENCE 418 AA; 47406 MW; 94905BA9B097F0DE CRC64;

Query Match 50.0%; Score 43; DB 11; Length 418;  
 Best Local Similarity 70.0%; Pred. No. 14;  
 Matches 7; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 NPWXXRGRI 10  
 |||||  
 Db 211 NPKWYDRGI 220

RESULT 8  
 09A382 PRELIMINARY; PRT; 433 AA.  
 AC 09A382;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PEPTIDOGLYCAN-BINDING PROTEIN, PUTATIVE.  
 GN CC3322.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 CC Caulobacter.  
 CX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hatt D.H.,  
 Kolony J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,  
 Rauterbach T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
 RT "Complete genome sequence of Caulobacter crescentus."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005994; AAK25284.1; -.  
 DR HSSP; P41052; 1LTM.  
 DR TIGR; CC3322; -.  
 DR InterPro: IPR002477; PG\_binding.  
 DR Pfam; PF04471; PG\_binding\_1; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

Query Match 48.8%; Score 42; DB 16; Length 433;  
 Best Local Similarity 60.0%; Pred. No. 22;  
 Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 2 PAWXXRGRI 11  
 |||||  
 Db 285 PAMWEARGVR 294

RESULT 9  
 018729 PRELIMINARY; PRT; 338 AA.  
 AC 018729;  
 DT 01-NOV-1996 (TREMBLrel. 01, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE C50C10.2 PROTEIN.  
 GN C50C10.2.  
 OS Caenorhabditis elegans.  
 CC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditidae;  
 OC Rhabditidae; Pelodermidae; Caenorhabditis.  
 CX NCBI\_TaxID=6239;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA McMurry A.A.;  
 RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.  
 RN [2]

RP SEQUENCE FROM N.A.  
 RX MEDLINE=99069613; PubMed=9851916;  
 RA none;  
 RT "Genome sequence of the nematode C.elegans: A platform for  
 RT investigating biology."  
 RL Science 282:2012-2018(1998).  
 DR EMBL; Z72505; CA96608.1; -.  
 DR InterPro: IPR003839; DUF215.  
 DR Pfam; PF02688; DUF215; 1.  
 SQ SEQUENCE 338 AA; 39053 MW; 1D05470746CB8543 CRC64;

Query Match 47.7%; Score 41; DB 5; Length 338;  
 Best Local Similarity 53.8%; Pred. No. 26;  
 Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 PAWXXRGIRPYG 14  
 |||||  
 Db 275 PYWQILFIRPIG 287

RESULT 10  
 09A7W7 PRELIMINARY; PRT; 545 AA.  
 AC 09A7W7;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPOTHETICAL PROTEIN CC1602.  
 GN CC1602.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 CC Caulobacter.  
 CX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B.,  
 Deboy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Hatt D.H.,  
 Kolony J.F., Smit J., Craven M.B., Khouli H., Shetty J., Berry K.,  
 Rauterbach T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O.,  
 Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.,  
 RT "Complete genome sequence of Caulobacter crescentus."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005835; AAK23581.1; -.  
 DR TIGR; CC1602; -.  
 KW Hypothetical protein: Complete proteome.  
 SQ SEQUENCE 545 AA; 60175 MW; FB3187D64F44C238 CRC64;

Query Match 47.7%; Score 41; DB 16; Length 545;  
 Best Local Similarity 46.7%; Pred. No. 43;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 2 PAWXXRGIRPYGRF 16  
 |||||  
 Db 242 PGTYRAAGRPYERY 256

RESULT 11  
 09T133 PRELIMINARY; PRT; 153 AA.  
 AC 09T133;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ENDOUNCLEASE.  
 GN 3.  
 OS Bacteriophage phlyeo3-12.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;

OC T7-like phages.  
 OX NCBI\_TaxID=110457;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pajunen M.I.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=2115554; PubMed=11222590;  
 RA Pajunen M.I., Kiljunen S.J., Soedergaard M.E.L., Skurnik M.;  
 RT "Complete genomic sequence of the lytic bacteriophage phiY603-12 of  
 RT *Yersinia enterocolitica* serotype O:3.";  
 RL J. Bacteriol. 183:1928-1937(2001).  
 DR EMBL: AJ251805; CAB63604.1; -  
 SO SEQUENCE 153 AA; 17640 MW; 211571BBDE6C641D CRC64;

Query Match 46.5%; Score 40; DB 9; Length 153;  
 Best Local Similarity 58.3%; Pred. No. 16;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 5 YXXRGIRPVGRF 16  
 DB 5 YAAAGVAKVCAF 16

RESULT 12  
 O86838 PRELIMINARY; PRT; 184 AA.  
 AC O86838;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DE PUTATIVE MEMBRANE PROTEIN.  
 GN SC9A10.05C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Harris D., Taylor K.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleiser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL031260; CAA20292.1; -  
 SO SEQUENCE 184 AA; 20178 MW; 58806A19F8FD6996 CRC64;

Query Match 46.5%; Score 40; DB 2; Length 184;  
 Best Local Similarity 50.0%; Pred. No. 20;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 3 AMYXXRGIRPVGRF 16  
 DB 90 AMVSAHGVRQVGF 103

RESULT 13  
 O92554

ID O92554 PRELIMINARY; PRT; 324 AA.  
 AC O92554;  
 DT 01-MAY-1999 (TREMBLrel. 10, Created)  
 DT 01-MAY-1999 (TREMBLrel. 10, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PUTATIVE ARAC FAMILY TRANSCRIPTIONAL REGULATOR.  
 GN SC9B2.05.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Saunders D.C., Harris D.;  
 RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleiser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinashi H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 CC SIMILARITY: BELONGS TO THE ARAC/XYLIS FAMILY OF TRANSCRIPTIONAL  
 CC REGULATORS.  
 DR EMBL: AL035212; CAA22785.1; -  
 DR InterPro: IPR000005; HTHArac.  
 DR Pfam: PF00165; HTH\_Arac; 1.  
 DR SMART: SM00342; HTH\_ARAC\_FAMILY\_1; 1.  
 DR PROSITE: PS00041; HTH\_ARAC\_FAMILY\_1; 1.  
 DR PROSITE: PS01124; HTH\_ARAC\_FAMILY\_2; 1.  
 KM DNA-binding; Transcription regulation.  
 SO SEQUENCE 324 AA; 34650 MW; 2AFB5C250A7D003D CRC64;

Query Match 46.5%; Score 40; DB 2; Length 324;  
 Best Local Similarity 57.1%; Pred. No. 37;  
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

OY 2 PAMYXXRGIRPVGR 15  
 DB 211 PAMYRALGDPVYGR 224

RESULT 14  
 O9PH18 PRELIMINARY; PRT; 390 AA.  
 AC O9PH18;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CYSTEINE SYNTHASE.  
 GN XFO128.  
 OS *Xylella fastidiosa*.  
 OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
 OC Xylella.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Agencio M.,  
 RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carver H.,  
 RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,

RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Frohne M., Furlan L.R.,  
 RA Gardier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitzina J.P.,  
 RA Krieger J.E., Kurume E.E., Laigret F., Lambais W.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marino C.L.,  
 RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montello-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA de Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Meldanis J., Setubal J.C.,  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003866; AAF82941.1; -.  
 DR InterPro: IPR001926; PALP.  
 DR Pfam: PF00291; PALP; 1.  
 DR Complete proteome.  
 SQ SEQUENCE 390 AA; 42479 MW; D294072E6A55188A CRC64;

Query Match 46.5%; Score 40; DB 16; Length 390;  
 Best Local Similarity 60.0%; Pred. No. 46;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 NPAWYXXRGI 10  
 DB 354 SPAMYAAGI 363

RESULT 15  
 ID 092DM2 PRELIMINARY; PRT; 486 AA.  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE LIN0791 PROTEIN.  
 GN LIN0791.  
 OS *Listeria innocua*.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Listeria.  
 OX NCBI\_TaxID:1642;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-CLIP 11262 / SEROVAR 6A;  
 RX PubMed-11679669;  
 RA Glaser P., Frangeul L., Buchrieser C., Rusniok C., Amend A.,  
 RA Baquero F., Berche P., Bloeker H., Brandt P., Chakraborty T.,  
 RA Charbit A., Chetoui F., Couve E., de Darvar A., Dehoux P.,  
 RA Domnan E., Dominguez-Bernal G., Duchud E., Durant L., Dussurget O.,  
 RA Ertlan K.D., Fsihl H., Garcia-del Portillo F., Garrido P.,  
 RA Gautier L., Goebel W., Gomez-Lopez N., Hain T., Hauf J., Jackson D.,  
 RA Jones L.-M., Kaerst U., Kreft J., Kuhn M., Kunst F., Kurapkai G.,  
 RA Madueno E., Maitournam A., Mata Vicente J., Ng E., Nedjari H.,  
 RA Nordisak G., Novella S., de Pablo B., Perez-Diaz J.-C., Purcell R.,  
 RA Remmel B., Rose M., Schlueter T., Simoes N., Tierrez A.,  
 RA Vazquez-Boland J.-A., Voss H., Wehlund J., Cossart P.,  
 RT "Comparative genomics of *Listeria species*,"  
 RL Science 294:849-852(2001).  
 DR EMBL: AL596166; CAC96023.1; -.  
 DR ListList: LIN00791; -.  
 DR Complete proteome.  
 SQ SEQUENCE 486 AA; 53127 MW; 1AEAA42D6F19ACB CRC64;

Query Match 46.5%; Score 40; DB 16; Length 486;  
 Best Local Similarity 50.0%; Pred. No. 58;  
 Matches 7; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 NPAWYXXRGI 14  
 DB 443 NPAWYXXRGI 456

Search completed: September 13, 2002, 09:29:20  
 Job time: 1065 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds

(without alignments)  
4.447 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_6\_21

Perfect score: 80

Sequence: 1 PAMVXXRGIRPYGRFX 16

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :  
1: A\_Geneseq\_032802.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77	96.2	15	21	AAV49293
2	77	96.2	20	18	AAW13394
3	77	96.2	20	20	AAW97236
4	77	96.2	20	21	AAW10365
5	77	96.2	20	21	AAV49294
6	77	96.2	20	22	AAW62534
7	77	96.2	20	22	AAW90992
8	77	96.2	21	18	AAW13395
9	77	96.2	21	21	AAW10366
10	77	96.2	21	22	AAW62535
11	77	96.2	22	18	AAW13396

12	77	96.2	22	21	AAW10367	Human oxytocin sec
13	77	96.2	22	22	AAW62536	Human CRH releasin
14	77	96.2	31	18	AAW13391	Human type G prote
15	77	96.2	31	20	AAW97235	Human type G prote
16	77	96.2	31	20	AAW97235	Human type G prote
17	77	96.2	31	21	AAW10362	Human type G prote
18	77	96.2	31	21	AAW10362	Human type G prote
19	77	96.2	31	21	AAW10362	Human type G prote
20	77	96.2	31	22	AAW90991	Human type G prote
21	77	96.2	31	22	AAW90995	Human type G prote
22	77	96.2	32	18	AAW13392	Human type G prote
23	77	96.2	32	21	AAW10363	Human type G prote
24	77	96.2	32	22	AAW62532	Human type G prote
25	77	96.2	33	18	AAW13393	Human type G prote
26	77	96.2	33	21	AAW10364	Human type G prote
27	77	96.2	33	22	AAW62533	Human type G prote
28	77	96.2	33	22	AAW62533	Human type G prote
29	77	96.2	33	22	AAW62533	Human type G prote
30	77	96.2	33	22	AAW62533	Human type G prote
31	77	96.2	33	22	AAW62533	Human type G prote
32	77	96.2	33	22	AAW62533	Human type G prote
33	77	96.2	33	22	AAW62533	Human type G prote
34	77	96.2	33	22	AAW62533	Human type G prote
35	77	96.2	33	22	AAW62533	Human type G prote
36	77	96.2	33	22	AAW62533	Human type G prote
37	77	96.2	33	22	AAW62533	Human type G prote
38	77	96.2	33	22	AAW62533	Human type G prote
39	77	96.2	33	22	AAW62533	Human type G prote
40	77	96.2	33	22	AAW62533	Human type G prote
41	77	96.2	33	22	AAW62533	Human type G prote
42	77	96.2	33	22	AAW62533	Human type G prote
43	77	96.2	33	22	AAW62533	Human type G prote
44	77	96.2	33	22	AAW62533	Human type G prote
45	77	96.2	33	22	AAW62533	Human type G prote

## ALIGNMENTS

RESULT 1	
ID AAV49293	standard; peptide: 15 AA.
AC AAV49293	
XX	
DT 22-FEB-2000	(first entry)
XX	
DE 19P2 ligand peptide fragment.	
XX	
KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;	
KW pituitary; regulatory mechanism; central nervous system; pancreatic.	
XX	
OS Homo sapiens.	
XX	
FM Key	Location/Qualifiers
FT Modified-site	15
FT	/note="C-terminal amide"
XX	
PN W09960112-A1.	
XX	
PD 25-NOV-1999.	
XX	
PF 20-MAY-1999;	99WO-JF02650.
XX	
PR 21-MAY-1998;	98JP-0140293.
XX	
PA (TAKE ) TAKEDA CHEM IND LTD.	
XX	
PI Matsumoto H, Kitada C, Hinuma S;	
XX	
DR WPI: 2000-039381/03.	
XX	
PT New monoclonal antibodies, useful in diagnosis, as drugs and in	

PF studying diseases related to ligand abnormality  
XX  
PS Disclosure; Page 26; 73pp; Japanese.  
XX  
CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AA19290-302 represent peptide fragments of the 19p2 ligand.  
CC  
XX  
SQ Sequence 15 AA:  
  
Query Match 96.2%; Score 77; DB 21; Length 15;  
Best Local Similarity 86.7%; Pred. No. 1.8e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 PAMVXXRGIRPYGRF 15  
||| |||||  
DB 1 pavyasrglrpygrf 15  
  
RESULT 2  
AAW31394  
ID AAW31394 standard; Peptide: 20 AA.  
XX  
AC AAW31394;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KM therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1996; 96JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02431.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a

CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreaticitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligolactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
CC  
XX  
SQ Sequence 20 AA:  
  
Query Match 96.2%; Score 77; DB 18; Length 20;  
Best Local Similarity 86.7%; Pred. No. 2.4e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 PAMVXXRGIRPYGRF 15  
||| |||||  
DB 6 pavyasrglrpygrf 20  
  
RESULT 3  
AAW97236  
ID AAW97236 standard; peptide: 20 AA.  
XX  
AC AAW97236;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Human type ligand polypeptide fragment.  
XX  
KM Rat type ligand; modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypovarianism; gonocyst coccogenesis;  
KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KM acromegaly; Charli-Frömmel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;  
KM contraceptive; placental function; choriocarcinoma; hydralid mole;  
KM interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KM abnormal lipidmetabolism; oxytocia.  
XX  
OS Homo sapiens.  
XX  
PN WO9658962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
PI WPI: 1999-105614/09.  
XX  
DR WPI: 1999-105614/09.  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 166; 241pp; English.  
XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for

CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Charli-Frömmel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 CC  
 SQ Sequence 20 AA;

Query Match 96.2%; Score 77; DB 20; Length 20;  
 Best Local Similarity 86.7%; Pred. NO. 2.4e-07;  
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PAMWXXRGIRPVGRF 15  
 |||| |||||  
 DB 6 pawyarsgirpvgrf 20

RESULT 4  
 AAB10365  
 ID AAB10365 standard; peptide: 20 AA.  
 AC AAB10365;  
 DT 24-NOV-2000 (first entry)  
 XX Human oxytocin secretion promoting peptide SEQ ID NO: 35.  
 DE  
 XX Human: oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KM caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.  
 XX  
 OS Homo sapiens.  
 XX  
 XX WO200038704-A1.  
 PN 06-JUL-2000.  
 PD  
 XX 22-DEC-1999; 99WO-JP07199.  
 PF  
 XX 25-DEC-1998; 98JP-0369585.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI WPI; 2000-452298/39.  
 DR  
 XX Physiologically-active polypeptide recognized as ligand by G  
 PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine  
 CC  
 XX Disclosure; Page 63; 72pp; Japanese.  
 PS  
 XX This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.

XX  
 SQ Sequence 20 AA;

Query Match 96.2%; Score 77; DB 21; Length 20;  
 Best Local Similarity 86.7%; Pred. NO. 2.4e-07;  
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PAMWXXRGIRPVGRF 15  
 |||| |||||  
 DB 6 pawyarsgirpvgrf 20

RESULT 5  
 AAY49294  
 ID AAY49294 standard; peptide: 20 AA.  
 AC AAY49294;  
 DT 22-FEB-2000 (first entry)  
 XX 19P2 ligand peptide fragment.  
 DE  
 XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Modified-site 20 /note="C-terminal amide"  
 FT  
 XX WO9960112-A1.  
 PN 25-NOV-1999.  
 PD  
 XX 20-MAY-1999; 99WO-JP02650.  
 PF  
 XX 21-MAY-1998; 98JP-0140293.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI WPI; 2000-039381/03.  
 DR  
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality  
 PT  
 XX Disclosure; Page 26; 73pp; Japanese.  
 PS  
 XX The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.  
 CC  
 XX Sequence 20 AA;

Query Match 96.2%; Score 77; DB 21; Length 20;  
 Best Local Similarity 86.7%; Pred. NO. 2.4e-07;  
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PAMWXXRGIRPVGRF 15  
 |||| |||||  
 DB 6 pawyarsgirpvgrf 20

RESULT 6  
AAG62534  
ID AAG62534 standard; peptide: 20 AA.  
XX  
AC AAG62534;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 35.  
XX  
XX  
KM Human: corticotrophin releasing hormone; CRH: G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KM Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
XX  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
XX  
DR WPI; 2001-355552/37.  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Claim 4; Page 75; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction, and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 20 AA:  
XX  
Query Match 96.2%; Score 77; DB 22; Length 20;  
Best Local Similarity 86.7%; Pred. No. 2.4e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 PAMYXXRGIRPVGRF 15  
DB 6 pawyaaarglrpvgrf 20  
IIII IIIIIIIIIII  
RESULT 7  
AAB90992  
ID AAB90992 standard; Peptide: 20 AA.  
XX  
AC AAB90992;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Prolactin releasing peptide SEQ ID NO:166.  
XX  
KM Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimide; maleimide group; amlno;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX

OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
XX  
PR 10-SEP-1999; 99US-0153406.  
XX  
PR 15-OCT-1999; 99US-0159783.  
XX  
PA (CONJ-) CONJUCHEM INC.  
XX  
PI Bidon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
PS Disclosure; Page 244; 733pp; English.  
XX  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 20 AA:  
XX  
Query Match 96.2%; Score 77; DB 22; Length 20;  
Best Local Similarity 86.7%; Pred. No. 2.4e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 PAMYXXRGIRPVGRF 15  
DB 6 pawyaaarglrpvgrf 20  
IIII IIIIIIIIIII  
RESULT 8  
AAW31395  
ID AAW31395 standard; Peptide: 21 AA.  
XX  
AC AAW31395;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 5.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX

PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI; 1997-363672/33.  
DR N-PSDB; AAV02432.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. A central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharide. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 21 AA:  
  
Query Match 96.2%; Score 77; DB 18; Length 21;  
Best Local Similarity 86.7%; Pred. No. 2.6e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 PAWYXXRGIRPVGRF 15  
      |||||  
DB 6 PAWYASRGIRPVGRF 20  
  
RESULT 9  
AAB10366  
ID AAB10366 standard; peptide; 21 AA.  
XX  
AC AAB10366;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
PN WPI; 2001-35552/37.  
XX

PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-452298/39.  
XX  
PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
PS Disclosure; Page 63; 72pp; Japanese.  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
SQ Sequence 21 AA:  
  
Query Match 96.2%; Score 77; DB 21; Length 21;  
Best Local Similarity 86.7%; Pred. No. 2.6e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 PAWYXXRGIRPVGRF 15  
      |||||  
DB 6 PAWYASRGIRPVGRF 20  
  
RESULT 10  
AAG62535  
ID AAG62535 standard; peptide; 21 AA.  
XX  
AC AAG62535;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 36.  
XX  
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
PN WPI; 2001-35552/37.  
XX

PT Use of G protein receptor ligand or peptide for controlling  
PT corticotropin releasing hormone secretion -  
XX  
PS Disclosure: Page 75; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 21 AA:  
  
Query Match 96.2%; Score 77; DB 22; Length 21;  
Best Local Similarity 86.7%; Pred. No. 2.6e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 PAMXXXXRGIRPVGRF 15  
||| |||||  
Db 6 pawyaasrgirpvgrf 20  
  
RESULT 11  
AAM31396  
ID AAM31396 standard; Peptide; 22 AA.  
XX  
AC AAM31396;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 6.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; propylactic;  
XX therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX WPI: 1997-363672/33.  
DR N-PSDB: AAV02433.  
XX  
XX Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX Claim 2; Page 186; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
CC sequence represented in AAM31396 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function

CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyperglycemia,  
CC hyperlipidemia, hypercholesterolemia, hyperglyceridemia,  
CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 22 AA:  
  
Query Match 96.2%; Score 77; DB 18; Length 22;  
Best Local Similarity 86.7%; Pred. No. 2.7e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 PAMXXXXRGIRPVGRF 15  
||| |||||  
Db 6 pawyaasrgirpvgrf 20  
  
RESULT 12  
AAB10367  
ID AAB10367 standard; peptide; 22 AA.  
XX  
AC AAB10367;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
XX veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX WPI: 2000-452298/39.  
DR  
XX  
PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
XX Disclosure: Page 64; 72pp; Japanese.  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This



RESULT 15  
 AAW97235  
 ID AAW97235 standard; peptide; 31 AA.  
 XX  
 AC AAW97235;  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE Human type ligand polypeptide fragment.  
 XX

KW Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; eulthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
 KW Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9858962-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 22-JUN-1998; 98WO-JP02765.  
 XX  
 PR 23-JUN-1997; 97JP-0165437.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 XX  
 DR MPI; 1999-105614/09.  
 XX

PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 secretion or placental function, e.g. for treating menopausal  
 syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 XX

PS Claim 3; Page 159; 241pp; English.  
 XX

CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
 CC syndrome, eulthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dysospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydralid mole, irruption mole, abortion, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 XX  
 SQ Sequence 31 AA;

Query Match 96.2%; Score 77; DB 20; Length 31;  
 Best Local Similarity 86.7%; Pred No. 3 9e-07;  
 Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PANYXXRGIRPVGRF 15  
 ||||| |||||  
 DB 17 pawyasrgilrvgf 31



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 seconds  
(without alignments)  
2.861 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_6\_21

Perfect score: 80

Sequence: 1 PAMXXXXRGIRPVGRFX 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/p/tdat2/2/1aa/5A.COMB.pep.\*

2: /cgn2\_6/p/tdat2/2/1aa/5B.COMB.pep.\*

3: /cgn2\_6/p/tdat2/2/1aa/6A.COMB.pep.\*

4: /cgn2\_6/p/tdat2/2/1aa/6B.COMB.pep.\*

5: /cgn2\_6/p/tdat2/2/1aa/PCtus.COMB.pep.\*

6: /cgn2\_6/p/tdat2/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	77	96.2	20	3	US-09-105-678A-46
2	77	96.2	20	4	US-08-776-971-64
3	77	96.2	20	4	US-09-421-208-46
4	77	96.2	21	3	US-09-105-678A-47
5	77	96.2	21	4	US-08-776-971-65
6	77	96.2	21	4	US-09-421-208-47
7	77	96.2	22	3	US-09-105-678A-48
8	77	96.2	22	4	US-08-776-971-66
9	77	96.2	22	4	US-09-421-208-48
10	77	96.2	31	3	US-09-105-678A-9
11	77	96.2	31	3	US-09-105-678A-43
12	77	96.2	31	4	US-08-776-971-61
13	77	96.2	31	4	US-09-421-208-9
14	77	96.2	31	4	US-09-421-208-43
15	77	96.2	32	3	US-09-105-678A-44
16	77	96.2	32	4	US-08-776-971-62
17	77	96.2	32	4	US-09-421-208-44
18	77	96.2	33	3	US-09-105-678A-45
19	77	96.2	33	4	US-08-776-971-63
20	77	96.2	33	4	US-09-421-208-45
21	77	96.2	87	4	US-08-776-971-59
22	77	96.2	87	4	US-08-776-971-135
23	77	96.2	87	4	US-08-776-971-138
24	77	96.2	20	3	US-09-105-678A-34
25	77	96.2	20	3	US-09-105-678A-40
26	77	96.2	20	4	US-08-776-971-8
27	77	96.2	20	4	US-08-776-971-50

28	76	95.0	20	4	US-08-776-971-98	Sequence 98, Appl
29	76	95.0	20	4	US-09-421-208-34	Sequence 34, Appl
30	76	95.0	20	4	US-09-421-208-40	Sequence 40, Appl
31	76	95.0	21	3	US-09-105-678A-35	Sequence 35, Appl
32	76	95.0	21	3	US-09-105-678A-41	Sequence 41, Appl
33	76	95.0	21	4	US-08-776-971-9	Sequence 9, Appl
34	76	95.0	21	4	US-08-776-971-51	Sequence 51, Appl
35	76	95.0	21	4	US-09-421-208-35	Sequence 35, Appl
36	76	95.0	21	4	US-09-421-208-41	Sequence 41, Appl
37	76	95.0	22	3	US-09-105-678A-36	Sequence 36, Appl
38	76	95.0	22	3	US-09-105-678A-42	Sequence 42, Appl
39	76	95.0	22	4	US-08-776-971-10	Sequence 10, Appl
40	76	95.0	22	4	US-08-776-971-52	Sequence 52, Appl
41	76	95.0	22	4	US-09-421-208-36	Sequence 36, Appl
42	76	95.0	22	4	US-09-421-208-42	Sequence 42, Appl
43	76	95.0	31	3	US-09-105-678A-7	Sequence 7, Appl
44	76	95.0	31	3	US-09-105-678A-8	Sequence 8, Appl
45	76	95.0	31	3	US-09-105-678A-31	Sequence 31, Appl

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suemaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 96.2% Score 77; DB 3; Length 20;  
Best Local Similarity 86.7%; Pred. No. 9.6e-08;  
Matches 13; Conservative 0; Mismatches 2; Indels 0;

OY 1 PAMYXXRGIRPVGRF 15  
DB 6 PAMYASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kikada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-FEB-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCJ/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 96.2%; Score 77; DB 4; Length 20;  
Best Local Similarity 86.7%; Pred. No. 9,6e-08;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PAMYXXRGIRPVGRF 15  
DB 6 PAMYASRGIRPVGRF 20

## RESULT 3

US-09-421-208-46  
Sequence 46, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 96.2%; Score 77; DB 4; Length 20;  
Best Local Similarity 86.7%; Pred. No. 9,6e-08;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 PAMYXXRGIRPVGRF 15  
DB 6 PAMYASRGIRPVGRF 20

## RESULT 4

US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 96.2%; Score 77; DB 3; Length 21;  
Best Local Similarity 86.7%; Pred. No. 1e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PAWYXRGRIPYGRF 15  
|||||  
Db 6 PAWYASRGIRPYGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habeta, Yugo  
Kawamata, Taji  
Hosoya, Masaaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 96.2%; Score 77; DB 4; Length 21;  
Best Local Similarity 86.7%; Pred. No. 1e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 PAWYXRGRIPYGRF 15  
|||||  
Db 6 PAWYASRGIRPYGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 1992 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 96.2%; Score 77; DB 4; Length 21;  
Best Local Similarity 86.7%; Pred. No. 1.e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVGRF 15  
|||||  
DB 6 PAMYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Motiyya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 96.2%; Score 77; DB 3; Length 22;  
Best Local Similarity 86.7%; Pred. No. 1.e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVGRF 15  
|||||  
DB 6 PAMYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSRO for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 96.2%; Score 77; DB 4; Length 22;  
Best Local Similarity 86.7%; Pred. No. 1.e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVGRF 15  
|||||  
DB 6 PAMYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

Sequence 48, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-48

Query Match 96.2%; Score 77; DB 4; Length 22;  
Best Local Similarity 86.7%; Pred. No. 1.1e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 PAWYXXRCIRPVGRF 15  
||| |||||  
Db 6 PAWYASRCIRPVGRF 20

RESULT 10  
US-09-105-678A-9  
Sequence 9, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-9

Query Match 96.2%; Score 77; DB 3; Length 31;  
Best Local Similarity 86.7%; Pred. No. 1.3e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
Oy 1 PAWYXXRCIRPVGRF 15  
||| |||||  
Db 17 PAWYASRCIRPVGRF 31

RESULT 11  
US-09-105-678A-43  
Sequence 43, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-43

Query Match 96.2%; Score 77; DB 3; Length 31;  
Best Local Similarity 86.7%; Pred. No. 1.5e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVGRF 15  
DB 17 PAMYASRGIRPVGRF 31

RESULT 12

US-08-776-971-61  
Sequence 61, Application US/08776971B  
Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji

TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997

CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 96.2%; Score 77; DB 4; Length 31;  
Best Local Similarity 86.7%; Pred. No. 1.5e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVGRF 15  
DB 17 PAMYASRGIRPVGRF 31

RESULT 13

US-09-421-208-9  
Sequence 9, Application US/09421208  
Patent No. 6258561

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
Moriya, Takeo  
Tanaka, Yoko

APPLICANT: Nishimura, Osamu

TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND

NUMBER OF SEQUENCES: 52

CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street

CITY: Boston

STATE: MA

COUNTRY: USA

ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/421,208  
FILING DATE:

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998

APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342

TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:

LENGTH: 31 amino acids  
TYPE: amino acid

STRANDEDNESS:  
TOPOLOGY: linear

MOLECULE TYPE: peptide  
US-09-421-208-9

QY 1 PAMYXXRGIRPVGRF 15  
DB 17 PAMYASRGIRPVGRF 31

Query Match 96.2%; Score 77; DB 4; Length 31;  
Best Local Similarity 86.7%; Pred. No. 1.5e-07;  
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVGRF 15  
DB 17 PAMYASRGIRPVGRF 31

RESULT 14

```

US-09-421-208-43
: Sequence 43, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA

:
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentln Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 43:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 31 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
:
: US-09-421-208-43
.

Query Match 96.2%; Score 77; DB 4; Length 31;
Best Local Similarity 86.7%; Pred. No. 1.5e-07;
Matches 13; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMXXXRGIRPYGRF 15
DB 17 PAMVASRGIRPYGRF 31

RESULT 15
US-09-105-678A-44
: Sequence 44, Application US/09105678A
: Patent No. 6103882
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
: APPLICANT: Moriya, Takeo
: APPLICANT: Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA

```

```

1 ZIP: 02109
2
3 COMPUTER READABLE FORM:
4
5 MEDIUM TYPE: Floppy disk
6
7 COMPUTER: IBM PC compatible
8
9 OPERATING SYSTEM: PC-DOS/MS-DOS
10
11 SOFTWARE: PatentIn Release #1.0, Version #1.30
12
13 CURRENT APPLICATION DATA:
14
15 APPLICATION NUMBER: US/09/105,678A
16
17 FILING DATE: 26-JUN-1998
18
19 PRIOR APPLICATION DATA:
20
21 APPLICATION NUMBER: JP 172118/1997
22
23 FILING DATE: 27-JUN-1997
24
25 ATTORNEY/AGENT INFORMATION:
26
27 NAME: Conlin, David G.
28
29 REGISTRATION NUMBER: 27,026
30
31 REFERENCE/DOCKET NUMBER: 48466-342
32
33 TELECOMMUNICATION INFORMATION:
34
35 TELEPHONE: 617-523-3400
36
37 TELEFAX: 617-523-6440
38
39 INFORMATION FOR SEQ ID NO: 44:
40
41 SEQUENCE CHARACTERISTICS:
42
43 LENGTH: 32 amino acids
44
45 TYPE: amino acid
46
47 STRANDEDNESS:
48
49 TOPOLOGY: linear
50
51 MOLECULE TYPE: peptide
52
53 US-09-105-678A-44

```

Query Match	Score 77;	DB 3;	Length 32;
Best Local Similarity	86.7%;	Pred. No. 1.6e-07;	
Matches 13;	Conservative 0;	Mismatches 2;	Indels 0;
QY	1	PAMXXXXRGIRPYGRF	15
Db	17	PAMVSRGIRPYGRF	31

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:00 ; Search time 172.41 Seconds  
(without alignments)  
8.917 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_6\_21

Perfect score: 80

Sequence: 1 PAMVXXRGIRPVGRF 16

Scoring table: BLOSUM62

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR\_71:\*

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	76	95.0	83	2 JC7607	prolactin-releasin
2	43	53.8	220	2 C83292	probable glutathio
3	43	53.8	333	2 H82852	hydroxybenzoate oc
4	43	53.8	719	2 S61046	ARPI protein - yea
5	42	52.5	433	2 H87660	peptidoglycan-bind
6	41	51.2	338	2 T20100	hypothetical prote
7	41	51.2	476	2 G64720	probable amino aci
8	41	51.2	476	2 G90629	probable inner mem
9	41	51.2	476	2 G85480	inner membrane tra
10	41	51.2	476	2 AG0502	probable amino-act
11	41	51.2	545	2 A87448	conserved hypotnet
12	40	50.0	184	2 T35841	probable membrane
13	40	50.0	324	2 T35901	probable aract faml
14	40	50.0	790	2 T47959	hypothetical prote
15	40	50.0	965	2 AE0418	valine--tRNA ligas
16	39	48.8	149	1 NRP937	endodeoxyribonucle
17	39	48.8	256	2 B95936	hypothetical prote
18	39	48.8	340	2 T20102	hypothetical prote
19	39	48.8	390	2 G82844	cysteine synthase
20	39	48.8	485	2 E83653	glutanyl-tRNA synt
21	39	48.8	953	2 E82068	valyl-tRNA synthet
22	38.5	48.1	877	1 IUBOON	N-cadherin precurs
23	38.5	48.1	906	1 IUBOON	N-cadherin 2 precurs
24	38.5	48.1	906	1 IUBOON	N-cadherin 2 precurs
25	38	47.5	127	2 B83157	hypothetical prote
26	38	47.5	158	2 F72725	probable ribosomal
27	38	47.5	194	2 D87357	hypothetical prote
28	38	47.5	238	2 H72646	hypothetical prote
29	38	47.5	359	2 T35179	oxidoreductase alp

30	38	47.5	398	2 B75254	acetate kinase - D
31	38	47.5	950	2 G83167	valyl-tRNA synthet
32	38	47.5	951	1 SYECVT	valine--tRNA ligas
33	38	47.5	951	2 C91283	valine tRNA synth
34	38	47.5	951	2 E86124	valine tRNA synth
35	38	47.5	2870	2 A35348	319k protein ndvB
36	37.5	46.9	154	2 AH0264	conserved hypotnet
37	37	46.2	105	2 A97505	hypothetical prote
38	37	46.2	105	2 AE2723	hypothetical prote
39	37	46.2	113	2 A12484	hypothetical prote
40	37	46.2	114	2 D33876	carcinoembryonic a
41	37	46.2	167	2 AB2796	acetyltransferase
42	37	46.2	167	2 B97575	hypothetical prote
43	37	46.2	178	2 D95026	ribosomal protein
44	37	46.2	178	2 D97897	50S ribosomal prot
45	37	46.2	293	2 C81408	hypothetical prote

## ALIGNMENTS

## RESULT 1

JC7607 prolactin-releasing peptide - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C:Accession: JC7607

R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, B.

Biochem. Biophys. Res. Commun. 281, 53-56, 2001

A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:

A:Reference number: JC7607; MUID:21092785; PMID:11178959

A:Contents: Spleen

A:Accession: JC7607

A:Molecule type: DNA

A:Residues: 183 <YAM>

A:Cross-references: DDBJ:AB040612; DDBJ:AB040613

C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior

release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PrRP

A:Introns: 33/1

Query Match 95.0%; Score 76; DB 2; Length 83;  
Best Local Similarity 86.7%; Pred. No. 2.2e-06;  
Matches 13; Conservative 0; Mismatches 2; Indels 0;

OY 1 PAMVXXRGIRPVGRF 15

Db 38 PAMVXXRGIRPVGRF 52

## RESULT 2

C83292 Probable glutathione S-transferase PA2821 [Imported] - Pseudomonas aeruginosa (strain

C:Species: Pseudomonas aeruginosa

C:Date: 15-Sep-2000 #sequence\_revision 15-Sep-2000 #text\_change 23-Mar-2001

C:Accession: C83292

R:Stover, C.K.; Pham, X.O.; Ertvin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.;

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Lardig, K.; L.

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pa

A:Reference number: A82950; MUID:20437337

A:Accession: C83292

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-220 <STO>

A:Cross-references: GB:AE004709; GB:AE004091; NID:g948904; PIDN:AG06209.1; GSPDB:GN

A:Experimental source: strain PA01

C:Genetics:

A:Gene: PA2821

C:Superfamily: plaiace glutathione transferase

Query Match 53.8%; Score 43; DB 2; Length 220;  
Best Local Similarity 64.3%; Pred. No. 3.9;  
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

QY 1 PAMYXXRGIRPVG 14  
||||| | | | | |

DB 39 PAMY--REISPLGR 50

RESULT 3  
H82852  
hydroxybenzoate octaprenyltransferase Xf0068 [Imported] - Xylella fastidiosa (strain 9a5  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: H82852  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82852  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <SIM>  
A:Cross-references: GB:AE003860; GB:AE003849; NID:g9104830; PIDN:AAF82881.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Relinack, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Birones, M.R.S.; Bueno, M.R.P.; Camargo, L.E.A.; Carraro, D.M.; Carrer, H  
as-Neto, E.; Docena, C.; El-Dorry, H.; Facinani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, B  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savaak  
M.; Teunhaio, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: Xf0068  
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 53.8%; Score 43; DB 2; Length 333;  
Best Local Similarity 61.5%; Pred. No. 5.8;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 PAMYXXRGIRPVG 13  
| | | | | | | | | |

DB 56 PYMKLRGDRPVG 68

RESULT 4  
S61046  
ARPI protein - yeast (Saccharomyces cerevisiae)  
N:Alternate names: protein D1478; protein YDL167c  
C:Species: Saccharomyces cerevisiae  
C:Date: 15-Feb-1996 #sequence\_revision 01-Mar-1996 #text\_change 29-Oct-1999  
C:Accession: S61046; S31139; S67719  
R:Pohl, T.M.  
submitted to the EMBL Data Library, November 1995  
A:Reference number: S61010  
A:Accession: S61046  
A:Molecule type: DNA  
A:Residues: 1-719 <POH>  
A:Cross-references: EMBL:267750; NID:g1061256; PIDN:CAA91579.1; PID:g1061272  
R:Wehner, E.P.; Rao, E.; Brendel, M.  
Mol. Gen. Genet. 237, 351-358, 1993  
A:Title: Molecular structure and genetic regulation of SFA, a gene responsible for resis

A:Reference number: S31138; MUID:93247548  
A:Accession: S31139  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-492; 'N', 494-719 <MEH>  
A:Cross-references: EMBL:X68020; NID:g577609; PIDN:CAA4159.1; PID:g288590  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992  
R:Pohl, T.M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67708  
A:Accession: S67719  
A:Molecule type: DNA  
A:Residues: 1-719 <POH>  
A:Cross-references: EMBL:274215; NID:g1431265; PIDN:CAA98741.1; PID:g253076; PID:g143  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:NRP1; ARPI  
A:Cross-references: MIPS:YDL167c; SGD:S0002326  
A:Map position: 4L

Query Match 53.8%; Score 43; DB 2; Length 719;  
Best Local Similarity 50.0%; Pred. No. 12;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 2 AMYXXRGIRPVG 13  
| | | | | | | | | |

DB 244 SMFTQYGVPRVG 255

RESULT 5  
H87660  
peptidoglycan-binding protein, probable [Imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C:Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 20-Apr-2001  
C:Accession: H87660  
R:Heiman, W.C.; Feldblum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg,  
B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gilm, M.L.; Haft, D.H.; Ko  
n, J.; Ermolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A:Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: A87249; MUID:21173698; PMID:11259647  
A:Accession: H87660  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-433 <STO>  
A:Cross-references: GB:AE005673; NID:g13425020; PIDN:AAK25284.1; GSPDB:GN00148  
C:Genetics:  
A:Gene: CC3322

Query Match 52.5%; Score 42; DB 2; Length 433;  
Best Local Similarity 60.0%; Pred. No. 11;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMYXXRGIR 10  
| | | | | | | | | |

DB 285 PAMWEARGVR 294

RESULT 6  
T20100  
hypothetical protein G50C10.2 - Caenorhabditis elegans  
C:Species: Caenorhabditis elegans  
C:Date: 15-Oct-1999 #sequence\_revision 15-Oct-1999 #text\_change 15-Oct-1999  
C:Accession: T20100  
R:McMurray, A.  
submitted to the EMBL Data Library, May 1996  
A:Reference number: Z19224  
A:Accession: T20100  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-338 <MIT>

A:Cross-references: EMBL:Z72505; PIDN:CAA96608.1; GSPDB:GN00023; CESP:C50C10.2  
A:Experimental source: clone C50C10  
C:Genetics:  
A:Gene: CESP:C50C10.2  
A:Map position: 5  
A:Introns: 74/3; 144/3; 267/3

Query Match 51.2%; Score 41; DB 2; Length 338;  
Best Local Similarity 53.8%; Pred. No. 13;  
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PAMVXXRCI 9  
Db 275 PAMVXXRCI 287

RESULT 7  
G64720  
probable amino acid transport protein yaaJ, sodium-dependent - Escherichia coli  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 26-Aug-1999  
C:Accession: G64720  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co-  
.A.; Rose, D.J.; Mau, B.; Shao, Y.  
S:Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: G64720  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-476 <BLAT>  
A:Cross-references: GB:AE000111; GB:U00096; NID:G1786181; PIDN:AMC73118.1; PID:G1786188;  
A:Experimental source: strain K-12, substrain MG1655  
C:Genetics:  
A:Gene: yaaJ  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein  
C:Keywords: amino acid transport; transmembrane protein  
F:10-26/Domain: transmembrane #status predicted <TM1>  
F:91-107/Domain: transmembrane #status predicted <TM2>  
F:142-158/Domain: transmembrane #status predicted <TM3>  
F:178-194/Domain: transmembrane #status predicted <TM4>  
F:208-224/Domain: transmembrane #status predicted <TM5>  
F:303-319/Domain: transmembrane #status predicted <TM6>  
F:349-365/Domain: transmembrane #status predicted <TM7>  
F:391-407/Domain: transmembrane #status predicted <TM8>  
F:414-430/Domain: transmembrane #status predicted <TM9>

Query Match 51.2%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 PAMVXXRCI 9  
Db 129 PAMVXXRCI 137

RESULT 8  
G90629  
probable inner membrane transport protein EGS0007 [imported] - Escherichia coli (strain  
C:Species: Escherichia coli  
C:Date: 18-Jul-2001 #sequence\_revision 18-Jul-2001 #text\_change 03-Aug-2001  
C:Accession: G90629  
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.  
gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H.  
DNA Res. 8, 11-22, 2001  
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and geno-  
A:Reference number: A99629; MUID:21156231; PMID:11258796  
A:Accession: G90629  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <HAY>

A:Cross-references: GB:BA000007; PIDN:BA833430.1; PID:G13359463; GSPDB:GN00154  
A:Experimental source: strain O157:H7, substrain RMD 0509952  
C:Genetics:  
A:Gene: EGS0007  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 51.2%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 PAMVXXRCI 9  
Db 129 PAMVXXRCI 137

RESULT 9  
G83480  
inner membrane transport protein [imported] - Escherichia coli (strain O157:H7, subst  
C:Species: Escherichia coli  
C:Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 14-Sep-2001  
C:Accession: G83480  
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; May  
hiller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apoda  
Nature 409, 529-533, 2001  
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.  
A:Reference number: A85480; MUID:21074935; PMID:11206551  
A:Accession: G83480  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <STO>  
A:Cross-references: GB:AE005174; NID:G12512682; PIDN:AMG54307.1; GSPDB:GN00145; UWGP:  
C:Experimental source: strain O157:H7, substrain EDL933  
C:Genetics:  
A:Gene: yaaJ  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 51.2%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 18;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Oy 1 PAMVXXRCI 9  
Db 129 PAMVXXRCI 137

RESULT 10  
AG0502  
probable amino-acid transport protein STY0006 [imported] - Salmonella enterica subsp.  
C:Species: Salmonella enterica subsp. enterica serovar Typh  
A:Note: this species has also been called Salmonella typh  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 27-Nov-2001  
C:Accession: AG0502  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Church  
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farr  
S.; Moule, S.; O'Gaora, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens,  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica se  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AG0502  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-476 <PAR>  
A:Cross-references: GB:AL513382; PIDN:CAD01159.1; PID:G16501289; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY0006  
C:Superfamily: sodium-dependent D-alanine/glycine transport protein

Query Match 51.2%; Score 41; DB 2; Length 476;  
Best Local Similarity 66.7%; Pred. No. 18;

Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 PAMYXXKGI 9  
 |||||  
 Db 129 PAMYMARGL 137

RESULT 11  
 A87448  
 conserved hypothetical protein CC1602 [Imported] - Caulobacter crescentus  
 C:Species: Caulobacter crescentus  
 C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
 C:Accession: A87448  
 R:Metzger, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
 B.; Laub, M.T.; Deboy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
 n, J.; Ermolova, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
 Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
 A>Title: Complete Genome Sequence of Caulobacter crescentus.  
 A:Reference number: A87249; MUID:21173698; PMID:11259647  
 A:Accession: A87448  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-545 <STO>  
 A:Cross-references: GB:AE005673; NID:g13422999; PIDN:AAK23581.1; GSPDB:GN00148  
 C:Genetics:  
 A:Gene: CC1602  
 C:Superfamily: hypothetical protein u1937b

Query Match 51.2%; Score 41; DB 2; Length 545;  
 Best Local Similarity 46.7%; Pred. No. 21;  
 Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PAMYXXKGI 15  
 |||||  
 Db 242 PGTYRAAGVPRV 256

RESULT 12  
 T35841  
 probable membrane protein - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T35841  
 R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, August 1998  
 A:Reference number: 221590  
 A:Accession: T35841  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-184 <HAR>  
 A:Cross-references: EMBL:AL031260; PIDN:CAA20292.1; GSPDB:GN00070; SCOEDB:SC9A10.05c  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC9A10.05c

Query Match 50.0%; Score 40; DB 2; Length 184;  
 Best Local Similarity 50.0%; Pred. No. 11;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

QY 2 AMYXXKGI 15  
 |||||  
 Db 90 AMYSAHGVGRV 103

RESULT 13  
 T35901  
 probable arac family transcription regulator - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
 C:Accession: T35901  
 R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.

submitted to the EMBL Data Library, January 1998  
 A:Reference number: 221593  
 A:Accession: T35901  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-324 <SAU>  
 A:Cross-references: EMBL:AL035212; PIDN:CAA22785.1; GSPDB:GN00070; SCOEDB:SC9B2.05  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC9B2.05

Query Match 50.0%; Score 40; DB 2; Length 324;  
 Best Local Similarity 57.1%; Pred. No. 19;  
 Matches 8; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

QY 1 PAMYXXKGI 14  
 |||||  
 Db 211 PAMYRALGDPVGR 224

RESULT 14  
 T47959  
 hypothetical protein F15G16.60 - Arabidopsis thaliana  
 C:Species: Arabidopsis thaliana (mouse-ear cross)  
 C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
 C:Accession: T47959  
 R:De Haan, M.; Maarse, A.C.; Grivelli, L.A.; Mewes, H.W.; Lemcke, K.; Mayer, K.F.X.; O  
 submitted to the Protein Sequence Database, January 2000  
 A:Reference number: Z24480  
 A:Accession: T47959  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-790 <DEH>  
 A:Cross-references: EMBL:AL132959  
 A:Experimental source: cultivar Columbia; BAC clone F15G16  
 C:Genetics:  
 A:Map position: 3  
 A:introns: 39/1; 678/2; 698/3; 773/2  
 A>Note: F15G16.60

Query Match 50.0%; Score 40; DB 2; Length 790;  
 Best Local Similarity 46.7%; Pred. No. 45;  
 Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

QY 1 PAMYXXKGI 15  
 |||||  
 Db 370 PRTYGRGLQPHGRW 384

RESULT 15  
 AE0418  
 valine--tRNA ligase (EC 6.1.1.9) [Imported] - Yersinia pestis (strain C092)  
 C:Species: Yersinia pestis  
 C>Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 27-Nov-2001  
 C:Accession: AE0418  
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.  
 deno-Taragda, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.  
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell  
 Nature 413, 523-527, 2001  
 A>Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
 A:Reference number: AB0001; MUID:21470413; PMID:11586360  
 A:Accession: AE0418  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-965 <KUR>  
 A:Cross-references: GB:AL590842; PIDN:CAC92673.1; PID:g15981369; GSPDB:GN00175  
 C:Genetics:  
 A:Gene: valS  
 C:Superfamily: valine--tRNA ligase  
 C:Keywords: ligase

Query Match 50.0%; Score 40; DB 2; Length 965;  
Best Local Similarity 57.1%; Pred. No. 55;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
OY 1 PAWYXXRCIRPVGR 14  
||||:||||  
DB 450 PAWIDCGNYYVGR 463

Search completed: September 13, 2002, 09:24:00  
Job time: 775 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:46 ; Search time 80.21 Seconds  
(without alignments)  
7.724 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_6\_21

Perfect score: 80

Sequence: 1 PAMVXXRGIRPVGRFX 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	77	96.2	87	1	PRRP_HUMAN
2	76	95.0	83	1	PRRP_RAT
3	76	95.0	98	1	PRRP_BOVIN
4	44	55.0	676	1	EXL1_HUMAN
5	43	53.8	719	1	NRP1_YEAST
6	42	52.5	402	1	EXL1_STRCO
7	41	51.2	476	1	YNAJ_ECOLI
8	39	48.8	149	1	EMRN_BPT7
9	39	48.8	485	1	SYE_BACHD
10	39	48.8	953	1	STV_VIBCH
11	38.5	48.1	877	1	CAD2_BOVIN
12	38.5	48.1	906	1	CAD2_HUMAN
13	38.5	48.1	906	1	CAD2_MOUSE
14	38.5	48.1	906	1	CAD2_RAT
15	38	47.5	158	1	RIL5_AERPE
16	38	47.5	637	1	MTWA_DROME
17	38	47.5	951	1	SYV_ECOLI
18	38	47.5	1723	1	AIM1_HUMAN
19	38	47.5	2832	1	NDVB_RHIME
20	37	46.2	179	1	RK6_GUTH
21	37	46.2	788	1	REC2_HAELN
22	36	45.0	329	1	Y493_MYCTU
23	36	45.0	994	1	SYV_XYLFA
24	35	43.8	383	1	CYCR_CHRYI
25	35	43.8	391	1	THIL_ZOORA
26	35	43.8	442	1	AROA_BORBR
27	35	43.8	442	1	AROA_BORPE
28	35	43.8	453	1	TBB2_GEOCN
29	35	43.8	460	1	SR54_HALNI
30	35	43.8	596	1	SDP_EIMBO
31	35	43.8	618	1	CIRI_CITFR
32	35	43.8	792	1	SYEB_XYLFA
33	35	43.8	922	1	GYRA_AERNSA

34	35	43.8	1187	1	TYK2_HUMAN	P29597	homo sapien
35	34.5	43.1	420	1	Z229_HUMAN	Q9UJW7	homo sapien
36	34.5	43.1	616	1	LEU1_CORGL	P42455	corynebacte
37	34.5	43.1	831	1	DPO1_THERL	P30313	thermus aqu
38	34.5	43.1	912	1	CAD2_CHICK	P10288	gallus gall
39	34	42.5	151	1	YAG6_RHIME	P13486	rhizodium m
40	34	42.5	224	1	YGCI_ECOLI	Q46898	escherichia
41	34	42.5	254	1	EMD_HUMAN	P50402	homo sapien
42	34	42.5	272	1	TRA2_DROYI	O02008	drosophila
43	34	42.5	306	1	R51_SYNPE	P46228	synechococc
44	34	42.5	328	1	R51A_SYNPE	P73530	synechocyst
45	34	42.5	344	1	RIL3_AERPE	O9YIM2	aeropyrum p

## ALIGNMENTS

RESULT ID	PRRP_HUMAN	STANDARD	PRP	87 AA.
AC	P81277			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Prolactin-releasing peptide precursor (PRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PRP31; Prolactin-releasing peptide PRP20].			
DE				
GN	PRH.			
OS	Homo sapiens (Human).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxID=9606;			
RP	[1]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RA	MEDLINE=96268781; PubMed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kikada C., Masuo Y., Asano T., Matsunoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RT	"A prolactin-releasing peptide in the brain.";			
RL	Nature 393:272-276(1998).			
RN	[2]			
RP	TISSUE SPECIFICITY.			
RX	PubMed=10498338;			
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Hinuma S., Sekiguchi M., Kikada C., Kurokawa T., Nishimura O., Onda H., Sunino Y., Fujino M.;			
RA	"Tissue distribution of prolactin-releasing peptide (PRP) and its receptor.";			
RL	Regul. Pept. 83:1-10(1999).			
CC	-1- FUNCTION: stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.			
CC	-1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
DR	EMBL: AB015419; BAA29027.1.			
DR	MIM: 602663.			
KW	Hormone; Amidation; Signal.			
FT	SIGNAL	1	22	
FT	PEPTIDE	23	53	BY SIMILARITY.
FT	PEPTIDE	34	53	PROLACTIN-RELEASING PEPTIDE PRP31.
FT	MOD RES	53	53	PROLACTIN-RELEASING PEPTIDE PRP20.
SO	SEQUENCE	87 AA.	9639 MW.	229A2F350CF9B1B CRC64; AMIDATION (G-54 PROVIDE AMIDE GROUP).







CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
ON NCBI\_TaxId=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=93(2);  
RA Seeger K.J., Harris D., Cerdano A.M., Parkhill J., Barrell B.G.,  
RL Submitt (Aug-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
CC -1- CATALYTIC ACTIVITY: Exonuclease cleavage in either 5' to 3'-  
CC or 3' to 5' direction to yield 5'-phosphomononucleotides.  
CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
CC SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE XSEB FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: AL391754; CAC05901.1; -  
DR InterPro: IPR003753; Exonuc\_VII\_L.  
DR InterPro: IPR002309; tRNA-synt\_2.  
DR Pfam: PF02601; Exonuc\_VII\_L.1.  
DR Pfam: PF01336; tRNA\_ant1.1.  
KW Hydrolyase; Nuclease; Exonuclease.  
SQ SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 52.5%; Score 42; DB 1; Length 402;  
Best Local Similarity 42.9%; Pred. No. 2.2;  
Matches 9; Conservative 1; Mismatches 3; Indels 8; Gaps 1;

OY 1 PAMYXXRG-----IRPVG 13  
DB 89 PEMYARQGLSLRAELIKPVG 109

RESULT 7  
YAAJ\_ECOLI STANDARD: PRT: 476 AA.  
ID YAAJ\_ECOLI  
AC P30143:  
DT 01-JUL-1993 (Rel. 26, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Putative transporter yaaJ.  
GN YAAJ OR B0007.  
OS Escherichia coli.  
OC Escherichia: Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.  
ON NCBI\_TaxId=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12;  
RA MEDLINE=92334977; PubMed=1630901;  
RX Yura T., Mori H., Nagai H., Nagata T., Ishihama A., Fujita N.,  
RA Isono K., Mizobuchi K., Nakata A.;  
RT "Systematic sequencing of the Escherichia coli genome: analysis of  
RT the 0-2.4 min region."  
RL Nucleic Acids Res. 20:3305-3308(1992).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -1- SIMILARITY: BELONGS TO THE SODIUM/ALANINE SYMPORTER FAMILY  
CC (SAP). STRONG, TO H. INFLUENZAЕ H10183.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL: D10483; -; NOT\_ANNOTATED\_CDS.  
DR EMBL: AE000111; AAC73118.1; -.  
DR Ecogene: EG11555; yaaJ.  
DR InterPro: IPR002293; AA\_rel\_permease\_1.  
DR InterPro: IPR001463; Na\_ala\_symp.  
DR Pfam: PF01235; Na\_Ala\_symp.1.  
DR PRINTS: PR00175; NAALASMPOR.  
DR PROSITE: PS00873; NA\_ALANINE\_SYMP.1.  
KW Hypothetical protein; Transmembrane; Inner membrane; Transport;  
KW Symport; Complete proteome.  
FT TRANSMEM 4 24 POTENTIAL.  
FT TRANSMEM 81 101 POTENTIAL.  
FT TRANSMEM 141 161 POTENTIAL.  
FT TRANSMEM 174 194 POTENTIAL.  
FT TRANSMEM 207 227 POTENTIAL.  
FT TRANSMEM 233 253 POTENTIAL.  
FT TRANSMEM 300 320 POTENTIAL.  
FT TRANSMEM 351 371 POTENTIAL.  
FT TRANSMEM 391 411 POTENTIAL.  
FT TRANSMEM 414 434 POTENTIAL.  
SQ SEQUENCE 476 AA; 51662 MW; 2F6EB2E12E126E63 CRC64;

Query Match 51.2%; Score 41; DB 1; Length 476;  
Best Local Similarity 66.7%; Pred. No. 3.9;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 PAMYXXRGI 9  
DB 129 PAMYMARGL 137

RESULT 8  
ENRN\_BP77 STANDARD: PRT: 149 AA.  
ID ENRN\_BP77  
AC P00641;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Endodeoxyribonuclease I (EC 3.1.21.2) (Endonuclease).  
GN 3.  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like phages  
ON NCBI\_TaxId=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=83241725; PubMed=6864790;  
RX Dunn J.J., Studier F.W.;  
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
RT locations of T7 genetic elements."  
RL J. Mol. Biol. 166:477-535(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC MEDLINE=82078034; PubMed=7310871;  
RX Dunn J.J., Studier F.W.;  
RT "Nucleotide sequence from the genetic left end of bacteriophage T7

```

RT DNA to the beginning of gene 4."
RL J. Mol. Biol. 148:303-330(1981).
CC -1- FUNCTION: ENDOEXOXYRIBONUCLEASE I, WHICH IS EXPRESSED IN THE LATE
CC STAGE, IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN
CC OF HOST DNA IN THE EARLY STAGE OF INFECTION. T7 DNA REPLICATES AS
CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA
CC LINEAR CONCATENERS SEVERAL GENOMES IN LENGTH. THE GENE 3 PRODUCT
CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE CONCATENERS.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphooligonucleotide end-products.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: V01127; CAA2345.1; -
DR EMBL: V01146; CAA24402.1; -
DR PIR: A00785; NEBP37.
DR PIR: S42301; S42301.
KM Hydrolyase; Nuclease; Endonuclease.
SQ SEQUENCE 149 AA; 17172 MW; D092AA2BE3743BC1 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 149;
Best Local Similarity 57.1%; Pred. No. 2.9;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 2 AMYXXRGIRPYGRF 15
   | : || | | |
Db 2 AGYCAKGIKRYGAF 15

RESULT 9
SYE_BACHD STANDARD; PRT; 485 AA.
ID SYE_BACHD
AC 09KGF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamy1-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
DE (GURS).
GN GUTX OR BH0109.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C-125 / JCM 9153;
RX MEDLINE=20512582; PubMed=11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis."
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC diphosphate + L-glutamyl-tRNA(Glu).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).

```

```

CC EMBL: AP001507; BAB03828.1; -
DR InterPro: IPR000924; tRNA-synt_1c.
DR InterPro: IPR001412; tRNA-synt_1i.
DR Pfam: PF00749; tRNA-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTGLU.
DR PROSITE: PS00178; AA-TRNA_LIGASE_1; 1.
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 11 21 "HIGH" REGION.
FT SITE 252 256 "KMSK" REGION.
FT BINDING 255 255 ATP (BY SIMILARITY).
SQ SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match 48.8%; Score 39; DB 1; Length 485;
Best Local Similarity 58.3%; Pred. No. 9.5;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 4 YXXRGIRPYGRF 15
   | : || | | |
Db 146 YEAKGIRPVAF 157

RESULT 10
SYV_VIBCH STANDARD; PRT; 953 AA.
ID SYV_VIBCH
AC 09KPF3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Valyl-tRNA synthetase (EC 6.1.1.9) (Valine--tRNA ligase) (VALRS).
DE VALS OR VC2503.
GN VALS OR VC2503.
OS Vibrio cholerae.
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.
OX NCBI_TaxID=666;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=EL TOR N16961 / SEROTYPE O1;
RX MEDLINE=20406833; PubMed=10952301;
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwinn M.L.,
RA Dodson R.J., Helt D.H., Hickey E.K., Peterson J.D., Umayam L.A.,
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,
RA McDonald L., Uitterback T., Fleischmann R.D., Nierman W.C., White O.,
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,
RA Fraser C.M.;
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio
RT cholerae."
RL Nature 406:477-483(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-valine + tRNA(Val) = AMP + diphosphate
CC + L-valyl-tRNA(Val).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL: AE004320; AA955645.1; -
DR HSSP: P96142; IGAX.
DR TIGR: VC2503; -
DR InterPro: IPR002300; tRNA-synt_1a.
DR InterPro: IPR001412; tRNA-synt_1i.
DR InterPro: IPR002303; tRNA-synt_val.
DR Pfam: PF00133; tRNA-synt_1i; 1.
DR PRINTS: PR00986; TRNASYNTVAL.
DR PROSITE: PS00186; AA-TRNA_LIGASE_1; 1.

```

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 42 52 "HIGH" REGION.  
 FT SITE 554 558 "KMSKS" REGION.  
 FT BINDING 557 557 ATP (BY SIMILARITY).  
 SO SEQUENCE 953 AA; 108170 MW; D93471A33CF4F69C CRC64;

Query Match 48.8%; Score 39; DB 1; Length 953;  
 Best Local Similarity 57.1%; Pred. No. 19;  
 Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 PAVYXXRGIRPVGR 14  
 ||||| : |||||  
 Db 436 PAVYDEGQNVGR 449

RESULT 11  
 CAD2\_BOVIN STANDARD; PRT; 877 AA.  
 AC P19534;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 01-FEB-1991 (Rel. 17, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neural-cadherin precursor (N-cadherin) (Cadherin-2) (Fragment).  
 GN CDH2.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovine; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90360979; PubMed=2390969;  
 RA Liao C.W., Cannon C., Power M.D., Kiboneka P.K., Rubin L.L.;  
 RT Identification and cloning of two species of cadherins in bovine  
 RT endothelial cells.";  
 RL EMBL J. 9:2701-2708(1990).

CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
 CC NEURONAL RECOGNITION MECHANISM.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-slb.ch/announce/>  
 CC or send an email to [license@isb-slb.ch](mailto:license@isb-slb.ch)).  
 CC -----

DR EMBL; X53615; CAA37677.1; -.  
 DR PIR; S11693; IJBON.  
 DR HSSP; P15116; INCI.  
 DR InterPro; IPR002126; Cadherin.  
 DR InterPro; IPR000233; Cadherin\_C-term.  
 DR Pfam; PF00028; cadherin; 5.  
 DR Pfam; PF01049; Cadherin\_C-term; 1.  
 DR PRINTS; PRO0205; CADHERIN.  
 DR SMART; SM00112; CA; 5.  
 DR PROSITE; PS00232; CADHERIN\_1; 3.  
 DR PROSITE; PS0268; CADHERIN\_2; 5.  
 KW Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat.  
 FT NON\_TER 1 1  
 FT PROPEP <1 130  
 FT CHAIN 131 877 NEURAL-CADHERIN.  
 FT DOMAIN 131 695 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 696 717 POTENTIAL.  
 FT DOMAIN 718 877 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 131 238 CADHERIN 1.

FT DOMAIN 239 363 CADHERIN 2.  
 FT DOMAIN 364 468 CADHERIN 3.  
 FT DOMAIN 469 574 CADHERIN 4.  
 FT DOMAIN 575 685 CADHERIN 5.  
 FT DOMAIN 834 849 SER-RICH.  
 FT CARBOHYD 161 161 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 244 244 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 296 296 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 373 373 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 543 543 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 663 663 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 SO SEQUENCE 877 AA; 96845 MW; 441B829ED871A249 CRC64;

Query Match 48.1%; Score 38.5; DB 1; Length 877;  
 Best Local Similarity 50.0%; Pred. No. 21;  
 Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

QY 1 PAV--YXXRGIRPVGR 15  
 ||| | | | |  
 Db 385 PAVNAIYRISGDPAGRF 402

RESULT 12  
 CAD2\_HUMAN STANDARD; PRT; 906 AA.  
 AC P19022; Q14923;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).  
 GN CDH2 OR CDHN OR NCAD.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=91016946; PubMed=2216790;  
 RA Reid R.A., Hemperly J.J.;  
 RT "Human N-cadherin: nucleotide and deduced amino acid sequence.";  
 RL Nucleic Acids Res. 18:5896-5896(1990).

CC [2]  
 CC REVISIONS TO 341; 699 AND 705.  
 CC [3]  
 CC Submitted (NOV-1990) to the EMBL/GenBank/DBJ databases.

CC SEQUENCE FROM N.A.  
 RX MEDLINE=92363956; PubMed=1500442;  
 RA Salomon D., Ayala O., Patel-King R., Hynes R.O., Gelger B.;  
 RT "Extrajunctional distribution of N-cadherin in cultured human  
 RT endothelial cells.";  
 RL J. Cell Sci. 102:7-17(1992).

CC [4]  
 CC SEQUENCE OF 160-906 FROM N.A.  
 RX MEDLINE=95047462; PubMed=2394753;  
 RA Walsh F.S., Barton C.H., Pitt W., Moore S.E., Kelsell D.,  
 RA Spurr N., Goodfellow P.N.;  
 RT "N-cadherin gene maps to human chromosome 18 and is not linked to the  
 RT E-cadherin gene.";  
 RL J. Neurochem. 55:805-812(1990).

CC [5]  
 CC SEQUENCE OF 1-20 FROM N.A.  
 RX MEDLINE=95048366; PubMed=7959764;  
 RA Wallis J.A., Fox M., Walsh F.S.;  
 RT "Structure of the human N-cadherin gene: YAC analysis and fine  
 RT chromosomal mapping to 18q11.2.";  
 RL Genomics 22:172-179(1994).  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THUS CONTRIBUTE TO THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN

CC NEURONAL RECOGNITION MECHANISM.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: X57548; CAA40773.1; -  
 DR EMBL: X54315; CAA38213.1; -  
 DR EMBL: S42303; AAB22854.1; -  
 DR EMBL: M34064; AAA03236.1; -  
 DR EMBL: 227420; CAA81799.1; -  
 DR PIR: A38870; ITHUCN.  
 DR HISSP: P15116; INCU.  
 DR MIM: 114020; -  
 DR InterPro: IPR002126; Cadherin.  
 DR InterPro: IPR000233; Cadherin\_C-term.  
 DR Pfam: PF000028; cadherin\_5.  
 DR Pfam: PF01049; cadherin\_C-term; 1.  
 DR PRINTS: PR00205; CADHERIN.  
 DR SMART: SM00112; CA; 5.  
 DR PROSITE: PS00232; CADHERIN\_1; 3.  
 DR PROSITE: PS50268; CADHERIN\_2; 5.  
 DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
 KM Signal.  
 FT SIGNAL. 1 23  
 FT PROPEP 24 159  
 FT CHAIN 160 906.  
 FT DOMAIN 160 724  
 FT TRANSSEM 725 746  
 FT DOMAIN 747 906  
 FT DOMAIN 160 267  
 FT DOMAIN 268 382  
 FT DOMAIN 383 497  
 FT DOMAIN 498 603  
 FT DOMAIN 604 714  
 FT DOMAIN 714 863  
 FT DOMAIN 863 878  
 FT CARBOHYD 190 190  
 FT CARBOHYD 273 273  
 FT CARBOHYD 325 325  
 FT CARBOHYD 402 402  
 FT CARBOHYD 572 572  
 FT CARBOHYD 651 651  
 FT CARBOHYD 692 692  
 FT CONFLICT 12 12  
 FT CONFLICT 16 16  
 FT CONFLICT 196 196  
 FT CONFLICT 212 212  
 FT CONFLICT 357 357  
 FT CONFLICT 867 867  
 FT SEQUENCE 906 AA; 99851 MW; 72DDC7BBB57C7A9C CRC64;  
 Query Match 48.1%; Score 38.5; DB 1; Length 906;  
 Best Local Similarity 50.0%; Pred. No. 22;  
 Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).  
 GN CDH2.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RX MEDLINE=69346748; PubMed=2762814;  
 RA Miyatani S., Shimamura K., Hata K., Nagafuchi A., Nose A.,  
 RA Matsunaga M., Hata K., Takeichi M.;  
 RT "Neural cadherin: role in selective cell-cell adhesion.";  
 RN Science 245:631-635(1989).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RA Tamura K.;  
 RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP PARTIAL SEQUENCE FROM N.A.  
 RC STRAIN-CS7BL/6;  
 RX MEDLINE=92409532; PubMed=1528849;  
 RA Miyatani S., Copeland N.G., Gilbert D.J., Jenkins N.A., Takeichi M.;  
 RT "Genomic structure and chromosomal mapping of the mouse N-cadherin  
 RT gene.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 89:8443-8447(1992).  
 [4]  
 RP DEVELOPMENTAL STAGE.  
 RC STRAIN-CS7BL/6; TISSUE-Testis;  
 RX MEDLINE=9703837; PubMed=8879495;  
 RA Munro S.B., Blaschuk O.W.;  
 RT "A comprehensive survey of the cadherins expressed in the testes of  
 RT fetal, immature, and adult mice utilizing the polymerase chain  
 RT reaction.";  
 RL Biol. Reprod. 55:822-827(1996).  
 [5]  
 RP X-RAY CRYSTALLOGRAPHY (1.9 ANGSTROMS) OF 160-267.  
 RX MEDLINE=95191680; PubMed=7885471;  
 RA Shapiro L., Fannon A.M., Kwong P.D., Thompson A., Lehmann M.S.,  
 RA Grubel G., Legrand J.-F., Als-Nielsen J., Colman D.R.,  
 RA Hendrickson W.A.;  
 RT "Structural basis of cell-cell adhesion by cadherins.";  
 RL Nature 374:327-337(1995).  
 [6]  
 RP X-RAY CRYSTALLOGRAPHY (3.4 ANGSTROMS) OF 160-374.  
 RX MEDLINE=98318235; PubMed=9655503;  
 RA Tamura K., Shan W.S., Hendrickson W.A., Colman D.R., Shapiro L.;  
 RT "Structure-function analysis of cell adhesion by neural (N-)  
 RT cadherin.";  
 RL Neuron 20:1153-1163(1998).  
 CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
 CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
 CC MANNER IN CONNECTING CELLS; CADHERINS MAY THIS CONTRIBUTE THE  
 CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
 CC NEURONAL RECOGNITION MECHANISM.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- DEVELOPMENTAL STAGE: EXPRESSED AT ALL STAGES OF TESTICULAR  
 CC DEVELOPMENT WITH HIGHEST LEVELS FOUND IN TESTES OF 21-DAY-OLD  
 CC MICE.  
 CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: M31131; AAA37353.1; -  
 DR EMBL: AB008811; BAA23549.1; -  
 DR EMBL: S45011; AAB23356.1; -  
 DR PIR: A32759; IUMSCN.

DR PDB; INCG; 10-JUL-95.  
DR PDB; INCH; 10-JUL-95.  
DR PDB; INCI; 10-JUL-95.  
DR PDB; INCI; 18-MAR-99.  
DR PDB; INCI; 18-MAR-99.  
DR MGI:88355; Cdh2.  
DR InterPro: IPR002126; Cadherin.  
DR InterPro: IPR002123; Cadherin\_C-term.  
DR Pfam: PF00028; cadherin; 5.  
DR Pfam: PF01049; Cadherin\_C-term; 1.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA; 5.  
DR PROSITE: PS00232; CADHERIN\_1; 3.  
DR PROSITE: PS00268; CADHERIN\_2; 5.  
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
KW Signal; 3D-structure.  
FT SIGNAL 1 23 POTENTIAL.  
FT PROPEP 24 159 NEURAL-CADHERIN.  
FT CHAIN 160 906 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 160 724 POTENTIAL.  
FT TRANSMEM 725 746 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 747 906 CADHERIN 1.  
FT DOMAIN 160 267 CADHERIN 2.  
FT DOMAIN 268 382 CADHERIN 3.  
FT DOMAIN 383 497 CADHERIN 4.  
FT DOMAIN 498 603 CADHERIN 5.  
FT DOMAIN 604 717 SER-RICH.  
FT DOMAIN 717 863 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 572 572 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 7 9 GRG -> APR (IN REF. 2).  
FT CONFLICT 565 565 Y -> N (IN REF. 2).  
FT CONFLICT 567 567 Q -> K (IN REF. 2).  
FT CONFLICT 624 624 A -> T (IN REF. 2).  
SQ SEQUENCE 906 AA; 99761 MW; 0206741B71707B7 CRC64;

Query Match 48.1%; Score 38.5; DB 1; Length 906;  
Best Local Similarity 50.0%; Pred. No. 22;  
Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

OY 1 PAM---YXXRGIRPVGRF 15  
DB 414 PAMNAVRISGDPPTGRF 431

RESULT 14  
CAD2\_RAT STANDARD; PRT; 906 AA.  
AC 0921Y3; 09R0T5;  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Neural-cadherin precursor (N-cadherin) (Cadherin-2).  
OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.  
RC STRAIN-SPRAGUE-DAWLEY; TISSUE-Testis;  
RX MEDLINE=98187820; PubMed=9528971;  
RA Chung S.S., Mo M.Y., Silvestrini B., Lee W.M., Cheng C.Y.;  
RT "Rat testicular N-cadherin: its complementary deoxyribonucleic acid  
cloning and regulation.";  
RL Endocrinology 139:1853-1862(1998).  
RN [2]  
RP SEQUENCE FROM N.A.

RC STRAIN-SPRAGUE DAWLEY; TISSUE-Brain;  
RA Asai K., Tada T., Yamamoto M., Tada A., Mizuno M., Elimoto T., Kato T.;  
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CADHERINS ARE CALCIUM DEPENDENT CELL ADHESION PROTEINS.  
CC THEY PREFERENTIALLY INTERACT WITH THEMSELVES IN A HOMOPHILIC  
CC MANNER IN CONNECTING CELLS. CADHERINS MAY THUS CONTRIBUTE TO THE  
CC SORTING OF HETEROGENEOUS CELL TYPES. N-CADHERIN MAY BE INVOLVED IN  
CC NEURONAL RECOGNITION MECHANISM.  
CC -1- SUBCELLULAR LOCATION: Type I membrane protein (By similarity).  
CC -1- TISSUE SPECIFICITY: IN TESTIS, EXPRESSED IN SERTOLI AND GERM  
CC CELLS.  
CC -1- SIMILARITY: CONTAINS 5 CADHERIN DOMAINS.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: AF097593; AAC83818.1; -;  
DR EMBL: AB017695; BAA84919.1; -;  
DR HSSP: P15116; INCI.  
DR InterPro: IPR002126; Cadherin.  
DR InterPro: IPR002123; Cadherin\_C-term.  
DR Pfam: PF00028; cadherin; 5.  
DR Pfam: PF01049; Cadherin\_C-term; 1.  
DR PRINTS: PR00205; CADHERIN.  
DR SMART: SM00112; CA; 5.  
DR PROSITE: PS00232; CADHERIN\_1; 3.  
DR PROSITE: PS00268; CADHERIN\_2; 5.  
DR Cell adhesion; Glycoprotein; Transmembrane; Calcium-binding; Repeat;  
KW Signal.  
FT SIGNAL 1 27 POTENTIAL.  
FT PROPEP 28 159 NEURAL-CADHERIN.  
FT CHAIN 160 906 EXTRACELLULAR (POTENTIAL).  
FT DOMAIN 160 724 POTENTIAL.  
FT TRANSMEM 725 745 CYTOPLASMIC (POTENTIAL).  
FT DOMAIN 746 906 CADHERIN 1.  
FT DOMAIN 160 267 CADHERIN 2.  
FT DOMAIN 268 382 CADHERIN 3.  
FT DOMAIN 383 487 CADHERIN 4.  
FT DOMAIN 498 603 CADHERIN 5.  
FT DOMAIN 604 717 SER-RICH.  
FT DOMAIN 717 863 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 190 190 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 273 273 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 325 325 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 402 402 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 572 572 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 622 622 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 651 651 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 692 692 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CONFLICT 7 9 G -> A (IN REF. 2).  
FT CONFLICT 48 48 T -> D (IN REF. 2).  
FT CONFLICT 153 153 A -> R (IN REF. 2).  
FT CONFLICT 646 646 A -> V (IN REF. 2).  
FT CONFLICT 658 658 N -> K (IN REF. 2).  
FT CONFLICT 724 724 T -> A (IN REF. 2).  
SQ SEQUENCE 906 AA; 99685 MW; 97D9937FD8D1F8B5 CRC64;

Query Match 48.1%; Score 38.5; DB 1; Length 906;  
Best Local Similarity 50.0%; Pred. No. 22;  
Matches 9; Conservative 0; Mismatches 6; Indels 3; Gaps 1;

OY 1 PAM---YXXRGIRPVGRF 15  
DB 414 PAMNAVRISGDPPTGRF 431

RESULT 15

```

RL15_AERPE
ID   RL15_AERPE          STANDARD:          PRT:   158 AA.
AC   Q9YF98;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   16-OCT-2001 (Rel. 40, Last annotation update)
DE   50S ribosomal protein L15P.
GN   RPL15P OR APE0343.
OS   Aeropyrum pernix.
OC   Archaea; Crenarchaeota; Desulfurococcaceae; Desulfurococcaceae;
OC   Aeropyrum.
OX   NCBI_TaxID=56636;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KI.
RX   MEDLINE=99310339; PubMed=10382966;
RA   Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Halkawa Y.,
RA   Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankaï A., Kosugi H.,
RA   Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H.,
RA   Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y.,
RA   Yamazaki J., Kushiida N., Oguchi A., Aoki K.-I., Kubota K.,
RA   Nakamura Y., Nomura N., Sako Y., Kikuchi H.;
RT   "Complete genome sequence of an aerobic hyper-thermophilic
RT   crenarchaeon, Aeropyrum pernix KI."
RL   DNA Res. 6:83-101(1999).
CC   -I- SIMILARITY: BELONGS TO THE L15P FAMILY OF RIBOSOMAL PROTEINS.
CC   -I-
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   -----
DR   EMBL: AP000059; BAA79298.1;
DR   InterPro: IPR001196; Ribosomal_L15.
DR   PROSITE: PS00475; RIBOSOMAL_L15; 1.
KW   Ribosomal protein; Complete proteome.
SQ   SEQUENCE 158 AA; 17480 MW; 01220322D4157321 CRC64;

Query Match          47.5%; Score 38; DB 1; Length 158;
Best Local Similarity 45.5%; Pred. No. 4.7;
Matches 5; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY   1 PAMXXXXRGIRP 11
    | | | | |
DB   54 PTVYGGKGFNP 64

```

Search completed: September 13, 2002, 09:30:46  
 Job time: 1136 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:20 ; Search time 311.85 Seconds  
(without alignments)  
8.876 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_6\_21  
Perfect score: 80  
Sequence: 1 PAMXXXXRGIRPVGRFX 16

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mmc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriopl:\*  
17: sp\_archaeop:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	67	83.8	117	13	Q9W624
2	46.5	58.1	420	2	Q931Z7
3	43	53.8	220	16	Q910Z2
4	43	53.8	333	16	Q9PH76
5	42	52.5	433	16	Q9A382
6	41	51.2	338	5	Q18729
7	41	51.2	545	16	Q9A7W7
8	40	50.0	153	9	Q9T133
9	40	50.0	184	2	Q86838
10	40	50.0	324	2	Q9Z554
11	40	50.0	340	10	Q91GZ0
12	40	50.0	562	5	Q9VN94
13	39	48.8	790	10	Q9M371
14	39	48.8	108	11	Q9D1V4
15	39	48.8	179	2	Q9K307
16	39	48.8	256	16	Q9ZVE9

17	39	48.8	340	5	Q18731	Q18731 caenorhabd1
18	39	48.8	390	16	Q9PH18	Q9PH18 xylella fas
19	39	48.8	425	16	Q986U6	Q986U6 rhizobium l
20	39	48.8	555	3	Q00050	Q00050 aspergillus
21	38.5	48.1	238	11	Q55075	Q55075 cricetus
22	38	47.5	97	2	Q33440	Q33440 pseudomonas
23	38	47.5	127	16	Q9HXA7	Q9HXA7 pseudomonas
24	38	47.5	194	16	Q9A9U0	Q9A9U0 caulobacter
25	38	47.5	238	17	Q9YEH0	Q9YEH0 aeropyrum p
26	38	47.5	359	2	Q88036	Q88036 streptomyce
27	38	47.5	398	16	Q9RR92	Q9RR92 deinococcus
28	38	47.5	416	16	Q99Z49	Q99Z49 streptococc
29	38	47.5	419	4	Q9Y276	Q9Y276 homo sapien
30	38	47.5	481	5	Q95WT7	Q95WT7 drosophila
31	38	47.5	536	5	Q95YN4	Q95YN4 drosophila
32	38	47.5	536	5	Q95PK7	Q95PK7 drosophila
33	38	47.5	536	5	Q95PK6	Q95PK6 drosophila
34	38	47.5	536	5	Q95PK5	Q95PK5 drosophila
35	38	47.5	536	5	Q95N07	Q95N07 drosophila
36	38	47.5	536	5	Q95NR6	Q95NR6 drosophila
37	38	47.5	536	5	Q95NR7	Q95NR7 drosophila
38	38	47.5	536	5	Q95NR0	Q95NR0 drosophila
39	38	47.5	703	5	Q95Z10	Q95Z10 trypanosoma
40	38	47.5	950	16	Q9HXH0	Q9HXH0 pseudomonas
41	37.5	46.9	333	2	Q9RJ10	Q9RJ10 streptomyce
42	37	46.2	114	11	Q63104	Q63104 rattus norv
43	37	46.2	160	10	Q9LMD3	Q9LMD3 oryza sativ
44	37	46.2	178	16	Q97SU7	Q97SU7 streptococc
45	37	46.2	234	11	Q63112	Q63112 rattus norv

## ALIGNMENTS

RESULT 1  
ID Q9W624 PRELIMINARY; PRT; 117 AA.  
AC Q9W624;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE C-RE AMIDE.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Satake H., Minakata H., Fujimoto M.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB020024; BAAV662.1; -  
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB2038C2B0 CRC64;

Query Match 83.8%; Score 67; DB 13; Length 117;  
Best Local Similarity 66.7%; Pred. No. 0.00014;  
Matches 10; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
QY 1 PAMXXXXRGIRPVGRFX 15  
DB 61 PAMVXXRGVRIRGRFX 75  
RESULT 2  
ID Q931Z7 PRELIMINARY; PRT; 420 AA.  
AC Q931Z7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
OS AUR2B.  
GN Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CCM3239;  
RA Kornanev J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CCM3239.";  
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AV033994; AAK61719.1; -  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 58.1%; Score 46.5; DB 2; Length 420;  
Best Local Similarity 52.6%; Pred. No. 2.8;  
Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

Oy 2 AMYXX-----RGIRPYGRF 15  
||| |||||  
Db 37 AMMAAVLRGSGIRPYGRF 55

RESULT 3  
ID 091022 PRELIMINARY: PRT: 220 AA.  
AC 091022:  
DT 01-MAR-2001 (TREMBlrel. 16, Created)  
DT 01-MAR-2001 (TREMBlrel. 16, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE PROBABLE GLUTATHIONE S-TRANSFERASE.  
GN PA2821.  
OS Pseudomonas aeruginosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Pseudomonadaceae;  
OC Pseudomonas.  
OX NCBI\_TaxID=287;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 15692 / PA01;  
RX MEDLINE=20437337; PubMed=10984043;  
RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
RA Hickey M.J., Brinkman F.S.L., Huftagle W.O., Kowalik D.J., Lagrou M.,  
RA Garner R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,  
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Laidig K., Lim R.M.,  
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,  
RA Reiter J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.,  
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an  
RT opportunistic pathogen.";  
RL Nature 406:959-964(2000).  
DR EMBL: AE004709; AAG06209.1; -  
DR InterPro: IPR004046; GST\_N.  
DR InterPro: IPR004045; GST\_N.  
KW Transferase; Complete proteome.  
SQ SEQUENCE 220 AA; 24716 MW; 6596183EACAA050 CRC64;

Query Match 53.8%; Score 43; DB 16; Length 220;  
Best Local Similarity 64.3%; Pred. No. 6;  
Matches 9; Conservative 1; Mismatches 2; Indels 2; Gaps 1;

Oy 1 PAMYXXRGIRPYGR 14  
||||| |||||  
Db 39 PAMY--REISPLGR 50

RESULT 4  
ID 09PH76 PRELIMINARY: PRT: 333 AA.  
AC 09PH76:  
DT 01-OCT-2000 (TREMBlrel. 15, Created)

DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYDROXYBENZONATE OCTAPRENYLTRANSFERASE.  
GN XE0068.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-9A5C.  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvaranga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bioness M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carer H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Fetto J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hobeisel J.D., Junqueira M.L., Kemper E.L., Klitzing J.P.,  
RA Krieger J.E., Kurumae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Queaglo R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Savaeski H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Melandis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).  
DR EMBL: AE003860; AAF82881.1; -  
DR InterPro: IPR000537; UblA.  
DR Pfam: PF01040; UblA; 1.  
DR PROSITE: PS00943; UblA; UNKNOWN\_1.  
KW Complete proteome.  
SQ SEQUENCE 333 AA; 37931 MW; ECF3FA716C962B95 CRC64;

Query Match 53.8%; Score 43; DB 16; Length 333;  
Best Local Similarity 61.5%; Pred. No. 9.4;  
Matches 8; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Oy 1 PAMYXXRGIRPYG 13  
||| |||||  
Db 56 PYWKIARGDRPYG 68

RESULT 5  
ID 09A382 PRELIMINARY: PRT: 433 AA.  
AC 09A382:  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE PEPTIDOLYCAN-BINDING PROTEIN, PUTATIVE.  
GN CC3322.  
OS Caulobacter crescentus.  
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
OC Caulobacter.  
OX NCBI\_TaxID=69394;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-ATCC 19089 / CB15;

```

RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Elesen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phake N.D., Ely B.,
RA Debroy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005994; AAK25284.1; -.
DR HSP: P41052; 1LTM.
DR TIGR: CC3322; -.
DR InterPro: IPR002477; PG_binding.
DR Pfam: PF01471; PG_binding_1; 1.
KW Complete proteome.
SQ SEQUENCE 433 AA; 46169 MW; F7DD18B9F4CA10A7 CRC64;

Query Match 52.5%; Score 42; DB 16; Length 433;
Best Local Similarity 60.0%; Pred. No. 19;
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 PAMXXRGIR 10
DB 285 PAMWEARGVR 294

RESULT 6
ID Q18729 PRELIMINARY; PRT; 338 AA.
AC Q18729;
DT 01-NOV-1996 (TReMBLrel. 01, Created)
DT 01-JAN-1998 (TReMBLrel. 05, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE C5OC10.2 PROTEIN.
GN C5OC10.2.
OS Caenorhabditis elegans.
OC Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC Rhabditidae; Peloderinae; Caenorhabditis.
OC NCBI_TaxID=6239;
RN [1]
RP SEQUENCE FROM N.A.
RA McNurray A.A.;
RL Submitted (MAY-1996) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-99069613; PubMed-9851916;
RA none;
RT "Genome sequence of the nematode C. elegans: A platform for
RT investigating biology.";
RL Science 282:2012-2018(1998).
DR EMBL: Z72505; CAA96608.1; -.
DR InterPro: IPR003839; DUF215.
DR Pfam: PF02688; DUF215; 1.
SQ SEQUENCE 338 AA; 39053 MW; ID05470746CB8543 CRC64;

Query Match 51.2%; Score 41; DB 5; Length 338;
Best Local Similarity 53.8%; Pred. No. 22;
Matches 7; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 PAMXXRGIRPVG 13
DB 275 PYWQILFIRPIG 287

RESULT 7
Q9A7W7 PRELIMINARY; PRT; 545 AA.
AC Q9A7W7;
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)

```

```

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HYPOTHETICAL PROTEIN CC1602.
GN CC1602.
OS Caulobacter crescentus.
OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;
OC Caulobacter.
OX NCBI_TaxID=69394;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19089 / CB15;
RX MEDLINE-21173698; PubMed-11259647;
RA Nierman W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,
RA Elesen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,
RA Potocka I., Nelson W.C., Newton A., Stephens C., Phake N.D., Ely B.,
RA Debroy R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H.,
RA Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K.,
RA Uterback T., Tran K., Wolf A., Vamathavan J., Ermolaeva M., White O.,
RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;
RT "Complete genome sequence of Caulobacter crescentus.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).
DR EMBL: AE005835; AAK23581.1; -.
DR TIGR: CC1602; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 345 AA; 60175 MW; FB3187D6F44C238 CRC64;

Query Match 51.2%; Score 41; DB 16; Length 545;
Best Local Similarity 46.7%; Pred. No. 37;
Matches 7; Conservative 2; Mismatches 6; Indels 0; Gaps 0;

QY 1 PAMXXRGIRPVG 15
DB 242 PGTYRAGVPRVRY 256

RESULT 8
ID Q9T133 PRELIMINARY; PRT; 153 AA.
AC Q9T133;
DT 01-MAY-2000 (TReMBLrel. 13, Created)
DT 01-MAY-2000 (TReMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE ENDONUCLEASE.
GN 3.
OS Bacteriophage phiyeo3-12.
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=110457;
RN [1]
RP SEQUENCE FROM N.A.
RA Pajunen M.I.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE-21125554; PubMed-11222590;
RA Pajunen M.I., Kiljunen S.J., Soederholm M.E.L., Skurnik M.;
RT "Complete genomic sequence of the lytic bacteriophage phiyeo3-12 of
RT Yersinia enterocolitica serotype O:3.";
RL J. Bacteriol. 183:1928-1937(2001).
DR EMBL: AJ251805; CAB63604.1; -.
SQ SEQUENCE 153 AA; 17640 MW; 211571BBDE6641D CRC64;

Query Match 50.0%; Score 40; DB 9; Length 153;
Best Local Similarity 58.3%; Pred. No. 14;
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 4 YXXRGIRPVG 15
DB 5 YAARGVRKVGAF 16

RESULT 9

```



09VN94  
ID 09VN94 PRELIMINARY: PRT: 562 AA.  
AC 09VN94;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-MAY-2000 (TREMBlrel. 13, Last annotation update)  
DE CG113 PROTEIN.  
GN CG113.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan D., Bouck J., Brokstein P., Brotlier P.,  
RA Burlis J.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunlov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fodor C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Hewland T.J., Hernandez J.R., Houck C.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ijigawa C.,  
RA Jaitai M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasbo P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mallet B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Mekulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclab J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster."  
RL Science 287:2185-2195(2000).  
DR EMBL; AE003603; AAF52054.1; -  
DR FLYBase; FBgn0037304; CG113.  
SQ SEQUENCE 562 AA; 63406 MW; AC09F5B6F5D2F970 CRC64;

Query Match 50.0%; Score 40; DB 5; Length 562;  
Best Local Similarity 40.0%; Pred. No. 58;  
Matches 6; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

OY 1 PAMXXRGIRPVGRF 15  
DB 52 PAMSSQMGVRSIAKF 66

RESULT 13  
O9M371  
ID O9M371 PRELIMINARY: PRT: 790 AA.  
AC O9M371;

DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE HYPOTHETICAL 87.4 KDA PROTEIN.  
GN F15G16.60.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA De Haan M., Maarse A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
RA Mayer K.F.X., Queller F., Salanoubat M.;  
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA EU Arabidopsis sequencing project;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL132959; CAB71097.1; -  
KW Hypothetical protein.  
SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 50.0%; Score 40; DB 10; Length 790;  
Best Local Similarity 46.7%; Pred. No. 83;  
Matches 7; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

OY 1 PAMXXRGIRPVGRF 15  
DB 370 PRTYSGRLQPHGRW 384

RESULT 14  
O9D1V4  
ID O9D1V4 PRELIMINARY: PRT: 108 AA.  
AC O9D1V4;  
DT 01-JUN-2001 (TREMBlrel. 17, Created)  
DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
DE CO30044P22R1K PROTEIN.  
GN CO30044P22R1K.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6J; TISSUE=CORPUS STRIATUM;  
RX MEDLINE=21085660; PubMed=11217851;  
RA Kawai J., Shinaagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
RA Arkawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
RA Atzawa K., Iwase M., Nishi K., Kiyosawa H., Kondo S., Yamanaka I.,  
RA Saito T., Okazaki Y., Gojobori T., Bono H., Kasuwa T., Saito R.,  
RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
RA Fleischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
RA Kuehl P., Lewis S., Matsuo Y., Nikaido I., Pesole G., Quackenbush J.,  
RA Schirli L.M., Staubli F., Suzuki R., Tomita M., Wagner L., Washio T.,  
RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barab G.,  
RA Blake J., Boffelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
RA Brownstein M.J., But C., Fletcher C., Fujita M., Gariboldi M.,  
RA Gustincich S., Hill D., Hofmann M., Hune D.A., Kamiya M., Lee N.H.,  
RA Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,  
RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
RA Sasaki H., Sato K., Schoenbach C., Seya T., Shibata Y., Storch K.-F.,  
RA Suzuki H., Toyooka K., Wang K.H., Weiss C., Whitaker C., Wilming L.,  
RA Wyszynski A., Yoshida K., Hasegawa Y., Kawai J., Kohsaki S.,  
RA Hayashizaki Y.;  
RT "Functional annotation of a full-length mouse cDNA collection."  
RL Nature 409:685-690(2001).  
DR EMBL; AK021149; BAB32306.1; -  
DR MGD; MGI:1924786; CO30044P22R1K.  
SQ SEQUENCE 108 AA; 11922 MW; 218DC6056B7112E0 CRC64;

Query Match 48.8%; Score 39; DB 11; Length 108;  
 Best Local Similarity 54.5%; Pred. No. 15;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 PAMXXXXGIRP 11  
 : : : : :  
 DB 38 PSWHNCRGISP 48

## RESULT 15

Q9K307 PRELIMINARY; PRT; 179 AA.  
 AC Q9K307;  
 DT 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
 DE PUTATIVE ACETYLTRANSFERASE.  
 GN 2SCG4.07C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Cerdeno A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Kinsahl H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 the 8 Mb streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL360034; CAB59578.1; -;  
 DR InterPro: IPR000182; Acetyltransf\_GCN5.  
 DR Pfam: PF00583; Acetyltransf; 1.  
 KW Transferase.  
 SO SEQUENCE 179 AA; 19784 MW; 0693898A177C64C8 CRC64;

Query Match 48.8%; Score 39; DB 2; Length 179;  
 Best Local Similarity 58.3%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;  
 QY 2 AMYXXRGIRPVG 13  
 : : : : :  
 DB 138 AMYERRGYRTG 149

Search completed: September 13, 2002, 09:29:20  
 Job time: 1065 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds  
(without alignments)  
4.169 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_7\_21  
Perfect score: 73  
Sequence: 1 AMYXXRGIRPGRFX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues  
Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: A\_Geneseq\_032802.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
23: /SIDSI/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	70	95.9	15	21	AAV49293	19P2 ligand peptid
2	70	95.9	15	21	AAV49296	19P2 ligand peptid
3	70	95.9	20	18	AAW31394	Human type G prote
4	70	95.9	20	20	AAW97236	Human type G prote
5	70	95.9	20	21	AAW10365	Human oxytocin sec
6	70	95.9	20	21	AAV49294	Human CRH releasin
7	70	95.9	20	22	AAW62534	Human CRH releasin
8	70	95.9	20	22	AAW90992	Proactin releasin
9	70	95.9	21	18	AAW31395	Human type G prote
10	70	95.9	21	21	AAW10366	Human oxytocin sec
11	70	95.9	21	22	AAW62535	Human CRH releasin

12	70	95.9	22	18	AAW31396	Human type G prote
13	70	95.9	22	21	AAW10367	Human oxytocin sec
14	70	95.9	22	21	AAW62536	Human CRH releasin
15	70	95.9	31	18	AAW31391	Human type G prote
16	70	95.9	31	20	AAW97235	Human type G prote
17	70	95.9	31	20	AAW67615	Human type G prote
18	70	95.9	31	21	AAW10362	Human oxytocin sec
19	70	95.9	31	21	AAV49291	Human CRH releasin
20	70	95.9	31	22	AAW62531	Human CRH releasin
21	70	95.9	31	22	AAW90991	Proactin releasin
22	70	95.9	31	22	AAW90995	Proactin releasin
23	70	95.9	32	18	AAW31392	Human type G prote
24	70	95.9	32	21	AAW10363	Human oxytocin sec
25	70	95.9	32	22	AAW62532	Human CRH releasin
26	70	95.9	33	18	AAW31393	Human type G prote
27	70	95.9	33	21	AAW10364	Human oxytocin sec
28	70	95.9	33	22	AAW62533	Human CRH releasin
29	70	95.9	87	18	AAW31390	Human type G prote
30	70	95.9	87	20	AAW97236	Human type G prote
31	70	95.9	87	21	AAW10361	Human oxytocin sec
32	70	95.9	87	22	AAW62530	Human CRH releasin
33	69	94.5	15	18	AAW31400	Synthetic ligand 1
34	69	94.5	15	20	AAW97230	C-terminal ligand
35	69	94.5	20	18	AAW31387	Rat type G protein
36	69	94.5	20	18	AAW31374	Bovine G protein-c
37	69	94.5	20	20	AAW97232	Bovine pituitary-d
38	69	94.5	20	20	AAW97234	Rat type ligand po
39	69	94.5	20	20	AAW95191	Bovine pituitary-d
40	69	94.5	20	20	AAW95175	Murine pituitary-d
41	69	94.5	20	21	AAW10350	Bovine oxytocin se
42	69	94.5	20	21	AAW10358	Rat oxytocin secre
43	69	94.5	20	21	AAV49301	19P2 ligand peptid
44	69	94.5	20	21	AAV49302	19P2 ligand peptid
45	69	94.5	20	22	AAW62519	Bovine CRH releas

## ALIGNMENTS

RESULT 1	AAV49293	standard; peptide: 15 AA.
ID	AAV49293	
AC	AAV49293	
XX		
DT	22-FEB-2000	(first entry)
XX		
DE	19P2 ligand peptide fragment.	
XX		
KW	Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;	
KW	pituitary; regulatory mechanism; central nervous system; pancreatic.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site	15
FT		/note="C-terminal amide"
XX		
PN	W09960112-A1.	
XX		
PD	25-NOV-1999.	
XX		
PF	20-MAY-1999;	99W0-JP02650.
XX		
PR	21-MAY-1998;	98JP-0140293.
XX		
PA	(TAKE ) TAKEDA CHEM IND LTD.	
XX		
PI	Matsumoto H, Kltada C, Hinuma S;	
XX		
DR	WPI: 2000-039381/03.	
XX		
PT	New monoclonal antibodies, useful in diagnosis, as drugs and in	

PT studying diseases related to ligand abnormality -

XX PS Disclosure; Page 26; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.

XX SQ Sequence 15 AA;

Query Match 95.9%; Score 70; DB 21; Length 15;  
 Best Local Similarity 85.7%; Pred. No. 1.6e-06;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
 ||| |||||||  
 Db 2 awyasrgirpvgrf 15

RESULT 2  
 AAY49296  
 ID AAY49296 standard; peptide; 15 AA.  
 XX AAY49296;  
 AC  
 XX 22-FEB-2000 (first entry)  
 DT  
 XX 19p2 ligand peptide fragment.  
 DE  
 XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15 /note="C-terminal amide"  
 FT  
 XX WO960112-A1.  
 PN 25-NOV-1999.  
 PD  
 XX 20-MAY-1999; 99WO-JP02650.  
 PF  
 XX 21-MAY-1998; 98JP-0140293.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI  
 XX WPI: 2000-039381/03.  
 DR  
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 PT  
 XX Disclosure; Page 27; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.

XX SQ Sequence 15 AA;

Query Match 95.9%; Score 70; DB 21; Length 15;  
 Best Local Similarity 85.7%; Pred. No. 1.6e-06;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
 ||| |||||||  
 Db 2 awyasrgirpvgrf 15

RESULT 3  
 AAW31394  
 ID AAW31394 standard; Peptide; 20 AA.  
 XX AAW31394;  
 AC  
 XX 06-APR-1998 (first entry)  
 DT  
 XX Human type G protein-coupled receptor ligand fragment 4.  
 DE  
 XX G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9724436-A2.  
 PD 10-JUL-1997.  
 XX  
 PF 26-DEC-1996; 96WO-JP03821.  
 PR  
 XX 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343371.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kitada C;  
 DR N-PSDB; AAW02431.  
 DR WPI: 1997-363672/33.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 PS  
 XX Claim 2; Page 185; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
 CC sequence represented in AAW31390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligodactylia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.



SO Sequence 20 AA;

Query Match 95.9%; Score 70; DB 18; Length 20;  
Best Local Similarity 85.7%; Pred. No. 2.1e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||  
Db 7 awyarsgllrpvgrf 20

RESULT 4  
AAW97236 standard; peptide: 20 AA.  
XX AAW97236;  
AC AAW97236;  
XX 06-MAY-1999 (first entry)  
XX  
XX Human type ligand polypeptide fragment.  
XX  
XX Rat type ligand; modulation; prolactin secretion;  
XX G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
XX menopausal syndrome; euthyroid; hypometabolism; lactation;  
XX pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
XX prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
XX acromegaly; Charli-Frommel syndrome; Argonz-del Castillo syndrome;  
XX Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
XX contraceptive; placental function; choriocarcinoma; hydatid mole;  
XX Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
XX abnormal lipidmetabolism; oxytocia.  
XX  
XX Homo sapiens.  
XX  
XX WC985862-A1.  
XX 30-DEC-1998.  
XX  
XX 22-JUN-1998; 98WO-JP02765.  
XX  
XX 23-JUN-1997; 97JP-0165437.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX WPI; 1999-105614/09.  
XX  
XX Use of G protein-coupled receptor ligands - for modulating prolactin  
XX secretion or placental function, e.g. for treating menopausal  
XX syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
XX Claim 3; Page 166; 241pp; English.  
XX  
XX The present sequence represents a human type ligand fragment. It  
XX is used in the course of the invention. The specification describes  
XX an agent for modulating prolactin secretion which comprises a  
XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
XX protein. The agents for promoting prolactin secretion can be used for  
XX treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
XX syndrome, euthyroid or hypometabolism. They can be used for promoting  
XX lactation in a domestic mammal and as an aphrodisiac. The agents for  
XX inhibiting prolactin secretion can be used for treating or preventing  
XX pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
XX prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
XX acromegaly, Charli-Frommel syndrome, Argonz-del Castillo syndrome,  
XX Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
XX The inhibitory agents can also be used as contraceptives. The agents for  
XX modulating placental function can be used for treating or preventing  
XX choriocarcinoma, hydatid mole, irruption mole, unthrifty fetus,  
XX abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX

SO Sequence 20 AA;

Query Match 95.9%; Score 70; DB 20; Length 20;  
Best Local Similarity 85.7%; Pred. No. 2.1e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||  
Db 7 awyarsgllrpvgrf 20

RESULT 5  
AAB10365 standard; peptide: 20 AA.  
XX AAB10365;  
AC AAB10365;  
XX 24-NOV-2000 (first entry)  
XX  
XX Human oxytocin secretion promoting peptide SEQ ID NO: 35.  
XX  
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
XX treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
XX caesarean section; artificial fertilization; galactostasis; goat; pig;  
XX veterinary medicine; milk production.  
XX  
XX Homo sapiens.  
XX  
XX WO200038704-A1.  
XX 06-JUL-2000.  
XX  
XX 22-DEC-1999; 99WO-JP07199.  
XX  
XX 25-DEC-1998; 98JP-0369585.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
XX WPI; 2000-452298/39.  
XX  
XX Physiologically-active polypeptide recognized as ligand by G  
XX protein-coupled receptor protein, for promoting secretion of oxytocin,  
XX as drugs for diseases relating to oxytocin secretion and in veterinary  
XX medicine -  
XX  
XX Disclosure; Page 63; 72pp; Japanese.  
XX  
XX This invention describes a novel oxytocin secretion-regulating agent  
XX which contains a ligand peptide or its salt for the G protein-coupled  
XX receptor protein. It is useful in the form of drugs for ameliorating,  
XX preventing and treating diseases relating to oxytocin secretion e.g.  
XX weak pains and atonic bleeding, before and after expulsion of placenta,  
XX uterine recovery failure, caesarean section, stoppage of artificial  
XX fertilization or galactostasis and is also applicable in veterinary  
XX medicine for promoting milk production in cow, goat and pig. This  
XX sequence represents a human peptide which acts as an oxytocin secretion  
XX promoter.  
XX  
XX Sequence 20 AA;  
XX

Query Match 95.9%; Score 70; DB 21; Length 20;  
Best Local Similarity 85.7%; Pred. No. 2.1e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||  
Db 7 awyarsgllrpvgrf 20

RESULT 6  
AA49294  
ID AAY49294 standard; peptide; 20 AA.  
XX  
AC AAY49294;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE 19P2 ligand peptide fragment.  
XX  
KM Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
XX pituitary; regulatory mechanism; central nervous system; pancreatic.  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 20  
FT /note="C-terminal amide"  
XX  
PN WO960112-A1.  
XX  
PD 25-NOV-1999.  
XX  
PF 20-MAY-1999; 99WO-JP02650.  
XX  
PR 21-MAY-1998; 98JP-0140293.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX WPI; 2000-039381/03.  
DR  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
XX studying diseases related to ligand abnormality -  
XX  
PS Disclosure; Page 26; 73pp; Japanese.  
XX  
CC The invention provides a monoclonal antibody which has a specific  
XX reaction with the part peptide of the C-terminal of 19P2 ligand or its  
XX derivative. The antibodies can be used in diagnosis or to treat or  
XX prevent diseases associated with abnormality in the pituitary function  
XX regulatory mechanism (e.g. promotion of prolactin secretion), central  
XX nervous regulatory mechanism, and pancreatic function regulatory  
XX mechanism. The antibody-based immunoassay can also be applied in  
XX clarifying the physiological functions of the ligand and its derivative.  
XX Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.  
SQ  
XX  
Sequence 20 AA;  
XX  
Query Match 95.9%; Score 70; DB 21; Length 20;  
Best Local Similarity 85.7%; Pred. No. 2.1e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AWYXXRGIRPVGRF 14  
Db ||| |||||||||  
7 awyasrgirpvgrf 20  
XX  
RESULT 7  
AAG62534  
ID AAG62534 standard; peptide; 20 AA.  
XX  
AC AAG62534;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 35.  
XX  
KM Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
XX analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
XX Addison's disease; adrenal gland hyperfunction; obesity.

XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
XX  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
XX WPI; 2001-355552/37.  
DR  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
XX corticotrophin releasing hormone secretion -  
XX  
PS Claim 4; Page 75; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
XX corticotrophin releasing hormone (CRH), involving the use of a G protein  
XX receptor ligand. This can be used to control the secretion of CRH and is  
XX useful as an analgesic or for treating, preventing or ameliorating  
XX diseases associated with CRH secretion such as hyperaldosteronism,  
XX hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
XX disease (including boredom, nausea, pigmentation, hypogonadism, hair  
XX loss, and hypotension), adrenal gland hypofunction and obesity. The  
XX present sequence is a peptide used in the exemplification of the  
XX invention.  
SQ  
XX  
Sequence 20 AA;  
XX  
Query Match 95.9%; Score 70; DB 22; Length 20;  
Best Local Similarity 85.7%; Pred. No. 2.1e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AWYXXRGIRPVGRF 14  
Db ||| |||||||||  
7 awyasrgirpvgrf 20  
XX  
RESULT 8  
AAB90992  
ID AAB90992 standard; peptide; 20 AA.  
XX  
AC AAB90992;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Prolactin releasing peptide SEQ ID NO:166.  
XX  
KM Protection; endogenous therapeutic peptide; peptidase; conjugation;  
XX blood component; modification; succinimidyl; maleimido group; amino;  
XX hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
XX  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
XX  
PR 10-SEP-1999; 99US-0153406.  
XX  
PR 15-OCT-1999; 99US-0159783.  
XX

PA (CONF-) CONNUCHEM INC.  
XX  
XX Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
XX  
PS Disclosure; Page 244; 733pp; English.  
XX  
XX The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimide and maleimide groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids,  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specifically as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
XX Sequence 20 AA:  
SQ

Query Match 95.9%; Score 70; DB 22; Length 20;  
Best Local Similarity 85.7%; Pred. No. 2.1e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||||  
Db 7 awyasrgirpvgrf 20

RESULT 9  
AAW31395  
ID AAW31395 standard; Peptide; 21 AA.  
XX  
XX AAW31395;  
AC  
XX 06-APR-1998 (first entry)  
DT  
XX Human type G protein-coupled receptor ligand fragment 5.  
DE  
XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9724436-A2.  
PN  
XX 10-JUL-1997.  
PD  
XX  
XX 26-DEC-1996; 96WO-JP03821.  
PE  
XX  
XX 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
PI Kawamata Y, Kitada C;

XX WPI; 1997-363672/33.  
DR N-PSDB; AAV02432.  
XX  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
XX This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polypnea,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
XX Sequence 21 AA:  
SQ

Query Match 95.9%; Score 70; DB 18; Length 21;  
Best Local Similarity 85.7%; Pred. No. 2.2e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||||  
Db 7 awyasrgirpvgrf 20

RESULT 10  
AAB10366  
ID AAB10366 standard; peptide; 21 AA.  
XX  
XX AAB10366;  
AC  
XX 24-NOV-2000 (first entry)  
DT  
XX Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
DE  
XX  
XX Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO200038704-A1.  
PN  
XX 06-JUL-2000.  
PD  
XX  
XX 22-DEC-1999; 99WO-JP07199.  
PE  
XX  
XX 25-DEC-1998; 98JP-0369585.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX Matsumoto H, Kitada C, Hinuma S;  
PI WPI; 2000-452298/39.  
XX  
XX Physiologically-active polypeptide recognized as ligand by G  
PT

PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PT as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine -  
 XX  
 XX  
 PS Disclosure: Page 63; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.

SO Sequence 21 AA;

Query Match 95.9%; Score 70; DB 21; Length 21;

Best Local Similarity 85.7%; Pred. No. 2.2e-06;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
 ||| ||||||||  
 Db 7 awyastgrlrvgrf 20

RESULT 11

AAAG62535  
 ID AAG62535 standard; peptide; 21 AA.

AC AAG62535;

DT 24-AUG-2001 (first entry)

DE Human CRH releasing protein related peptide SEQ ID NO: 36.

KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
 KW Addison's disease; adrenal gland hyperfunction; obesity.

OS Homo sapiens.

PN WO200135984-A1.

PD 25-MAY-2001.

PF 17-NOV-2000; 2000WO-JP08119.

PR 18-NOV-1999; 99JP-0327900.

PR 26-SEP-2000; 2000JP-0297073.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Kitada C, Matsumoto H, Hinuma S;

DR WPI: 2001-35552/37.

PT Use of G protein receptor ligand or peptide for controlling  
 PT corticotrophin releasing hormone secretion -

PS Disclosure: Page 75; 90pp; Japanese.

CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The

CC present sequence is a peptide used in the exemplification of the  
 CC invention.

SO Sequence 21 AA;

Query Match 95.9%; Score 70; DB 22; Length 21;

Best Local Similarity 85.7%; Pred. No. 2.2e-06;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
 ||| ||||||||  
 Db 7 awyastgrlrvgrf 20

RESULT 12

AAAW31396  
 ID AAW31396 standard; peptide; 22 AA.

AC AAW31396;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 6.

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prolactin;  
 KW therapeutic agent.

OS Homo sapiens.

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kitada C;

DR WPI: 1997-363672/33.

DR N-PSDB: AAV02433.

PS Claim 2; Page 186; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
 CC sequence represented in AAW31396 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyperglycemia, hyperlipidemia,  
 CC hypercholesterolemia, hypercholesterolemia, hyperglycemia, hyperlipidemia,  
 CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligogalactia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting

CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 22 AA;

Query Match 95.9%; Score 70; DB 18; Length 22;  
Best Local Similarity 85.7%; Pred. No. 2.4e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPYGRF 14  
||| |||||||||  
Db 7 awyaarglrpygrf 20

RESULT 13  
AAB10367  
ID AAB10367 standard; peptide: 22 AA.  
XX  
AC AAB10367;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
XX

KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
XX

OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX

PD 06-JUL-2000.  
XX  
PE 22-DEC-1999; 99WO-JP07199.  
XX

PR 25-DEC-1998; 98JP-0369585.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Matsumoto H, Kitada C, Hinuma S;  
XX

DR WPI; 2000-452298/39.  
XX

PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
XX  
XX

PS Disclosure; Page 64; 72pp; Japanese.  
XX

CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
XX

SQ Sequence 22 AA;

Query Match 95.9%; Score 70; DB 21; Length 22;  
Best Local Similarity 85.7%; Pred. No. 2.4e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPYGRF 14  
||| |||||||||  
Db 7 awyaarglrpygrf 20

RESULT 14  
AAG62536  
ID AAG62536 standard; peptide: 22 AA.  
XX  
AC AAG62536;  
XX

DT 24-AUG-2001 (first entry)  
XX

DE Human CRH releasing protein related peptide SEQ ID NO: 37.  
XX

KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX

OS Homo sapiens.  
XX

PN WO200135984-A1.  
XX

PD 25-MAY-2001.  
XX

PE 17-NOV-2000; 2000WO-JP08119.  
XX

PR 18-NOV-1999; 99JP-0327900.  
XX

PR 26-SEP-2000; 2000JP-0297073.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Kitada C, Matsumoto H, Hinuma S;  
XX

DR WPI; 2001-355552/37.  
XX

PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
XX

PS Disclosure; Page 75; 90pp; Japanese.  
XX

CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hyperfunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
XX

SQ Sequence 22 AA;

Query Match 95.9%; Score 70; DB 22; Length 22;  
Best Local Similarity 85.7%; Pred. No. 2.4e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPYGRF 14  
||| |||||||||  
Db 7 awyaarglrpygrf 20

RESULT 15

AAW31391  
ID AAW31391 standard; Peptide: 31 AA.  
XX

AC AAW31391;  
XX

DT 06-APR-1998 (first entry)  
XX

DE Human type G protein-coupled receptor ligand fragment 1.  
XX

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;

KW therapeutic agent.  
 XX  
 OS Homo sapiens.  
 XX  
 PN MO9724436-A2.  
 XX  
 PD 10-JUL-1997.  
 XX  
 PF 26-DEC-1996; 96WO-JP03821.  
 XX  
 PR 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343371.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kltada C;  
 XX  
 DR WPI: 1997-363672/33.  
 DR N-PSDB: AAV02428.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2: Page 184; 258pp; English.  
 XX  
 CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
 CC sequence represented in AAW21390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
 CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX

Sequence 31 AA:

Query Match 95.9%; Score 70; DB 18; Length 31;  
 Best Local Similarity 85.7%; Pred. No. 3.4e-06;  
 Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
 QY 1 AWXXRGIRPVGRF 14  
 ||| |||||  
 DB 18 awyargirpvgrf 31

Search completed: September 13, 2002, 09:18:35  
 Job time: 500 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 seconds  
(without alignments)  
2.682 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_7\_21

Perfect score: 73

Sequence: 1 AMYXXRGIRPVGRFX 15

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

1: /cgn2\_6/prodata/2/1aa/5A\_COMB.pep:\*

2: /cgn2\_6/prodata/2/1aa/5B\_COMB.pep:\*

3: /cgn2\_6/prodata/2/1aa/6A\_COMB.pep:\*

4: /cgn2\_6/prodata/2/1aa/6B\_COMB.pep:\*

5: /cgn2\_6/prodata/2/1aa/PCTUS\_COMB.pep:\*

6: /cgn2\_6/prodata/2/1aa/backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	70	95.9	20	3	US-09-105-678A-46
2	70	95.9	20	4	US-08-776-971-64
3	70	95.9	20	4	US-09-421-208-46
4	70	95.9	21	3	US-09-105-678A-47
5	70	95.9	21	4	US-08-776-971-65
6	70	95.9	22	4	US-09-421-208-47
7	70	95.9	22	4	US-09-105-678A-48
8	70	95.9	22	4	US-08-776-971-66
9	70	95.9	22	4	US-09-421-208-48
10	70	95.9	31	3	US-09-105-678A-9
11	70	95.9	31	4	US-09-105-678A-43
12	70	95.9	31	4	US-08-776-971-61
13	70	95.9	31	4	US-09-421-208-9
14	70	95.9	31	4	US-09-105-678A-43
15	70	95.9	32	3	US-09-105-678A-44
16	70	95.9	32	4	US-08-776-971-62
17	70	95.9	32	4	US-09-421-208-44
18	70	95.9	33	3	US-09-105-678A-45
19	70	95.9	33	4	US-08-776-971-63
20	70	95.9	33	4	US-09-421-208-45
21	70	95.9	87	4	US-08-776-971-59
22	70	95.9	87	4	US-08-776-971-135
23	70	95.9	87	4	US-08-776-971-138
24	69	94.5	15	4	US-08-776-971-93
25	69	94.5	20	3	US-09-105-678A-34
26	69	94.5	20	3	US-09-105-678A-40
27	69	94.5	20	4	US-08-776-971-8

28	69	94.5	20	4	US-08-776-971-50	Sequence 50, Appl
29	69	94.5	20	4	US-08-776-971-98	Sequence 98, Appl
30	69	94.5	20	4	US-09-421-208-34	Sequence 34, Appl
31	69	94.5	20	4	US-09-421-208-40	Sequence 40, Appl
32	69	94.5	21	3	US-09-105-678A-35	Sequence 35, Appl
33	69	94.5	21	3	US-09-105-678A-41	Sequence 41, Appl
34	69	94.5	21	4	US-08-776-971-9	Sequence 9, Appl1
35	69	94.5	21	4	US-08-776-971-51	Sequence 51, Appl
36	69	94.5	21	4	US-09-421-208-35	Sequence 35, Appl
37	69	94.5	21	4	US-09-421-208-41	Sequence 41, Appl
38	69	94.5	22	3	US-09-105-678A-36	Sequence 36, Appl
39	69	94.5	22	3	US-09-105-678A-42	Sequence 42, Appl
40	69	94.5	22	4	US-08-776-971-10	Sequence 10, Appl
41	69	94.5	22	4	US-08-776-971-52	Sequence 52, Appl
42	69	94.5	22	4	US-09-421-208-36	Sequence 36, Appl
43	69	94.5	22	4	US-09-421-208-42	Sequence 42, Appl
44	69	94.5	31	3	US-09-105-678A-7	Sequence 7, Appl1
45	69	94.5	31	3	US-09-105-678A-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
; Sequence 46, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIORITY APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-46

Query Match 95.9%; Score 70; DB 3; Length 20;  
Best Local Similarity 85.7%; Pred. No. 1.2e-06;  
Matches 12; Conservative 0; Mismatches 2; Gaps 0;

OY 1 AMYXRGIRPVGRF 14  
DB 7 AMYASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64  
; Sequence 64, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: Fastseq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

## Query Match

Best Local Similarity 95.9%; Score 70; DB 4; Length 20;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXRGIRPVGRF 14  
DB 7 AMYASRGIRPVGRF 20

## RESULT 3

US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-46

## Query Match

Best Local Similarity 95.9%; Score 70; DB 4; Length 20;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXRGIRPVGRF 14  
DB 7 AMYASRGIRPVGRF 20

## RESULT 4

US-09-105-678A-47  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston



STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 95.9%; Score 70; DB 3; Length 21;  
Best Local Similarity 85.7%; Pred. No. 1.3e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||  
DB 7 AMYASRGIRPVGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/087/6971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kltada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: Internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 95.9%; Score 70; DB 4; Length 21;  
Best Local Similarity 85.7%; Pred. No. 1.3e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
||| |||||  
DB 7 AMYASRGIRPVGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 95.9%; Score 70; DB 4; Length 21;  
Best Local Similarity 85.7%; Pred. No. 1.3e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14  
||| |||||  
DB 7 AMYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
; Sequence 48, Application US/09105678A  
; Patent No. 6103882

GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 95.9%; Score 70; DB 3; Length 22;  
Best Local Similarity 85.7%; Pred. No. 1.3e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14  
||| |||||  
DB 7 AMYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
; Sequence 66, Application US/08776971B  
; Patent No. 6228984

GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hinata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kikada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 95.9%; Score 70; DB 4; Length 22;  
Best Local Similarity 85.7%; Pred. No. 1.3e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14  
||| |||||  
DB 7 AMYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

```
; Sequence 48, Application US/09421208
; Patent No. 6258561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-48

Query Match          95.9%; Score 70; DB 4; Length 22;
Best Local Similarity 85.7%; Pred. No. 1.3e-06;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 AMYXXRGIRPVGRF 14
   ||| |||||
Db 7 AMYASRGIRPVGRF 20
```

```
RESULT 10
US-09-105-678A-9
; Sequence 9, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-9
```

```
Query Match          95.9%; Score 70; DB 3; Length 31;
Best Local Similarity 85.7%; Pred. No. 1.9e-06;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY 1 AMYXXRGIRPVGRF 14
   ||| |||||
Db 18 AMYASRGIRPVGRF 31
```

```
RESULT 11
US-09-105-678A-43
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESSES:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
```

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-105-678A-43  
MOLECULE TYPE: peptide

Query Match  
Best Local Similarity 95.9%; Score 70; DB 3; Length 31;  
Best Local Similarity 85.7%; Pred. No. 1.9e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14  
DB 18 AMYASRGIRPVGRF 31

RESULT 12  
US-08-776-971-61  
Sequence 61, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Sho-ji  
Kitade, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03621  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match  
Best Local Similarity 95.9%; Score 70; DB 4; Length 31;  
Best Local Similarity 85.7%; Pred. No. 1.9e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14  
DB 18 AMYASRGIRPVGRF 31

RESULT 13  
US-09-421-208-9  
Sequence 9, Application US/09421208  
Patent No. 6238561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
US-09-421-208-9  
MOLECULE TYPE: peptide

Query Match  
Best Local Similarity 95.9%; Score 70; DB 4; Length 31;  
Best Local Similarity 85.7%; Pred. No. 1.9e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14  
DB 18 AMYASRGIRPVGRF 31

RESULT 14

US-09-421-208-43  
; Sequence 43, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-43

Query Match 95.9%; Score 70; DB 4; Length 31;  
Best Local Similarity 85.7%; Pred. No. 1.9e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPYGRF 14  
||| |||||||||  
Db 18 AMYASRGIRPYGRF 31

RESULT 15  
US-09-105-678A-44  
; Sequence 44, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-44

Query Match 95.9%; Score 70; DB 3; Length 32;  
Best Local Similarity 85.7%; Pred. No. 2e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPYGRF 14  
||| |||||||||  
Db 18 AMYASRGIRPYGRF 31

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:00 ; Search time 172.41 Seconds  
(without alignments)  
8.360 Million cell updates/sec

Title: US-09-446-543a-73-COPY\_7\_21  
Perfect score: 73  
Sequence: 1 AMYXXRGIRPVGRFX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	69	94.5	83	2	JC7607
2	43	58.9	719	2	S61046
3	40	54.8	184	2	T35841
4	39	53.4	149	1	NEBP37
5	39	53.4	236	2	E95936
6	39	53.4	485	2	E83663
7	38	52.1	194	2	D87357
8	38	52.1	238	2	H72646
9	38	52.1	333	2	H82852
10	38	52.1	398	2	B75254
11	38	52.1	2870	2	A35348
12	37.5	51.4	154	2	AH0264
13	37	50.7	105	2	A97505
14	37	50.7	105	2	A97505
15	37	50.7	113	2	A12484
16	37	50.7	114	2	D33876
17	37	50.7	293	2	C81408
18	37	50.7	293	2	G97409
19	37	50.7	293	2	AH2627
20	37	50.7	320	2	G83340
21	37	50.7	338	2	D95851
22	37	50.7	343	2	E95986
23	37	50.7	395	2	T34570
24	37	50.7	769	2	T51112
25	37	50.7	788	2	A63045
26	36.5	50.0	472	2	A83331
27	36	49.3	142	2	S60747
28	36	49.3	220	2	C83292
29	36	49.3	276	2	D70817

30	36	49.3	302	2	B70631	hypothetical prote
31	36	49.3	338	2	T20100	hypothetical prote
32	36	49.3	488	2	B89957	hypothetical prote
33	36	49.3	503	2	A82193	sun/nucleolar prot
34	36	49.3	545	2	A87448	conserved hypothet
35	36	49.3	600	2	D84258	alcohol dehydrogen
36	36	49.3	790	2	T47959	probable oxidoredu
37	36	49.3	966	2	S43878	hypothetical prote
38	36	49.3	975	2	T30816	trwC protein - Esc
39	36	49.3	1067	2	T28653	macrophage colony-
40	36	49.3	68	2	D72428	hypothetical prote
41	35	47.9	105	2	D81319	hypothetical prote
42	35	47.9	115	2	G72674	hypothetical prote
43	35	47.9	132	2	D82954	hypothetical prote
44	35	47.9	178	2	D95026	ribosomal protein
45	35	47.9				

## ALIGNMENTS

```
RESULT 1
JC7607
Proactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya,
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat proactin-releasing peptide gene:
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:Introns: 33/1

Query Match 94.5%; Score 69; DB 2; Length 83;
Best Local Similarity 85.7%; Pred. No. 1.2e-05;
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPVGRF 14
Db 39 AMYXXRGIRPVGRF 52

RESULT 2
S61046
ARPI protein - yeast (Saccharomyces cerevisiae)
N:Alternate names: protein D1478; protein YDL167c
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999
C:Accession: S61046; S31139; S67719
R:Poll, T.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61046
A:Accession: S61046
A:Molecule type: DNA
A:Residues: 1-719 <POH>
A:Cross-references: EMBL:Z67750; NID:91061256; PIDN:CAA91579.1; PID:91061272
R:Werner, E.P.; Rao, E.; Brendel, M.
Mol. Gen. Genet. 237, 351-358, 1993
A:Title: Molecular structure and genetic regulation of SFA, a gene responsible for re
A:Reference number: S31138; MUID:93247548
A:Accession: S31139
A:Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
```

A:Residues: 1-492, 'N', 494-719 <MEH>  
A:Cross-references: EMBL:X68020; NID:g577609; PIDN:CAA48159.1; PID:g288590  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992  
R:Polh, T.M.  
Submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67708  
A:Accession: S67719  
A:Molecule type: DNA  
A:Residues: 1-719 <POW>  
A:Cross-references: EMBL:274215; NID:g1431265; PIDN:CAA98741.1; PID:e253076; PID:g143126  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:NRP1: ARP1  
A:Cross-references: MIPS:YDL167c; SGD:S0002326  
A:Map position: 4L

Query Match 58.9%; Score 43; DB 2; Length 719;  
Best Local Similarity 50.0%; Pred. No. 6.6;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

Qy 1 AMYXXRGIRPVG 12  
11: 111111  
Db 244 SMFTGYGVRRVG 255

RESULT 3  
T35841  
probable membrane protein - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C>Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T35841  
R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
Submitted to the EMBL Data Library, August 1998  
A:Reference number: Z21590  
A:Accession: T35841  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-184 <HAR>  
A:Cross-references: EMBL:AL031260; PIDN:CAA20292.1; GSPDB:GN00070; SCOEDB:SC9A10.05c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC9A10.05c

Query Match 54.8%; Score 40; DB 2; Length 184;  
Best Local Similarity 50.0%; Pred. No. 6.2;  
Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

Qy 1 AMYXXRGIRPVGRF 14  
11: 111111  
Db 90 AWVSAGVROVGRF 103

RESULT 4  
NEBP37  
endodeoxyribonuclease I (EC 3.1.21.-) - phage T7  
C:Species: phage T7  
C>Date: 01-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 11-Jun-1999  
C:Accession: B94615; C92866; S42301; A00785  
R:Dunn, J.J.; Thompson, K.  
Submitted to the Nucleic Acid Sequence Database, September 1982  
A:Reference number: A94615  
A:Accession: B94615  
A:Molecule type: DNA  
A:Residues: 1-149 <DU1>  
R:Dunn, J.J.; Studier, F.W.  
J. Mol. Biol. 148, 303-330, 1981  
A:Title: Nucleotide sequence from the genetic left end of bacteriophage T7 DNA to the be  
A:Reference number: A92866; MUID:82078034  
A:Accession: C92866  
A:Molecule type: DNA  
A:Residues: 1-149 <DU2>

A:Cross-references: GB:V01127; NID:g15498; PIDN:CAA24345.1; PID:g15517  
R:Dunn, J.J.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7  
A:Reference number: S42283; MUID:83341725  
A:Accession: S42301  
A:Molecule type: DNA  
A:Residues: 1-149 <DUN>  
A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24402.1; PID:g15581  
A:Note: the authors did not translate the codon for residue 1  
C:Comment: Endodeoxyribonuclease I, which is expressed in the late stage, is necessary  
s a linear monomer. In the late stage, the T7 DNA replicates via linear concatamers s  
C:Genetics:  
A:Gene: 3  
A:Map position: 25.64-26.76  
C:Superfamily: phage T7 exodeoxyribonuclease I  
C:Keywords: hydrolase

Query Match 53.4%; Score 39; DB 1; Length 149;  
Best Local Similarity 57.1%; Pred. No. 7.7;  
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

Qy 1 AMYXXRGIRPVGRF 14  
11: 111111  
Db 2 AGYGAKGIRRVGAF 15

RESULT 5  
E95936  
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) magaplasmid ps  
C:Species: Sinorhizobium meliloti  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: E95936  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholzer, F.J.; Her  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing e  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: E95936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-256 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAC49157.1; PID:g15140642; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid pSymb  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubl  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau  
hebaull, P.; Vandenbol, M.; Vorholzer, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMD21253  
A:genome: plasmid

Query Match 53.4%; Score 39; DB 2; Length 256;  
Best Local Similarity 54.5%; Pred. No. 13;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

Qy 2 WYXXRGIRPVG 12  
11: 111111  
Db 185 WHGTRCGRPVG 195

RESULT 6  
E83663  
glutanyl-tRNA synthetase gtx [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: E83663





A:Experimental source: strain R1  
C:Genetics:  
A:Gene: DR2602  
A:Map position: 1  
C:Superfamily: acetate kinase

Query Match 52.1%; Score 38; DB 2; Length 398;  
Best Local Similarity 50.0%; Pred. No. 31;  
Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
QY 1 AMYXXRGIRPVG 12  
DB 161 AMYXHGVRKY 172

RESULT 11  
A35548  
319k protein ndvb - Rhizobium meliloti  
C:Species: Rhizobium meliloti  
C:Date: 21-Sep-1990 #sequence\_revision 21-Sep-1990 #text\_change 08-Oct-1999  
C:Accession: A35548  
R:Lepl, L.; Dylen, T.; Ditta, G.S.; Hellnakl, D.R.; Stanfield, S.W.  
J. Biol. Chem. 265, 2843-2851, 1990  
A:Title: The ndvb locus of Rhizobium meliloti encodes a 319-kDa protein involved in the  
A:Reference number: A35548; MUID:90153914  
A:Accession: A35548  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2870 <IEU>  
A:Cross-references: GB:J05219; NID:g152270; PIDN:AAA26305.1; PID:g152271  
C:Keywords: transmembrane protein

Query Match 52.1%; Score 38; DB 2; Length 2870;  
Best Local Similarity 50.0%; Pred. No. 2,1e+02;  
Matches 5; Conservative 2; Mismatches \*3; Indels 0; Gaps 0;

QY 2 WYXXRGIRPV 11  
DB 1126 WYRTGLEPM 1135

RESULT 12  
AH0264  
conserved hypothetical protein YPO2172 [imported] - Yersinia pestis (strain CO92)  
C:Species: Yersinia pestis  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C:Accession: AH0264  
R:Parhill, J.; Wren, B.W.; Thomson, N.R.; Tibball, R.W.; Holden, M.T.G.; Prentice, M.B.;  
deno-Farraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrell,  
Nature 413, 523-527, 2001  
A:Title: Genome sequence of Yersinia pestis, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AH0264  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-154 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC90980.1; PID:g15980175; GSPDB:GN00175  
C:Genetics:  
A:Gene: YPO2172  
C:Superfamily: hypothetical protein HI0277

Query Match 51.4%; Score 37.5; DB 2; Length 154;  
Best Local Similarity 61.5%; Pred. No. 15;  
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;  
QY 2 WYXXRGIRPVGR 13  
DB 124 WYIIDGIRPGLGR 136

RESULT 13  
A97505  
hypothetical protein AGR\_C\_2196 [imported] - Agrobacterium tumefaciens (strain C58, C  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: A97505  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Quirillo, B.; Goldm  
A.; Liu, F.; Wollam, C.; Allinger, M.; Dougherty, D.; Scott, C.; Lappas, C.; Markelz,  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; PMID:11743194  
A:Accession: A97505  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-105 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK86994.1; PID:g15156234; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_2196  
A:Map position: circular chromosome

Query Match 50.7%; Score 37; DB 2; Length 105;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 AMYXXRGIR 9  
DB 54 AMYAAAGVQ 62

RESULT 14  
AE2723  
hypothetical protein Atu1191 [imported] - Agrobacterium tumefaciens (strain C58, Dupo  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AE2723  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo  
erage, G.; Gillet, W.; Grant, C.; Genthner, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl  
; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kam  
ster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AE2723  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-105 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AAL42203.1; PID:g17739595; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu1191  
A:Map position: circular chromosome

Query Match 50.7%; Score 37; DB 2; Length 105;  
Best Local Similarity 55.6%; Pred. No. 13;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIR 9  
DB 54 AMYAAAGVQ 62

RESULT 15  
AI2484  
hypothetical protein alr7057 [imported] - Anabaena sp. (strain PCC 7120) plasmid PCC7  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AI2484  
R:Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Irigu

Nakazaki, N.; Shimo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
 DNA Res. 8, 205-213, 2001  
 A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
 A:Reference number: AB1807; MUID:21595285; PMID:11759840  
 A:Accession: AI2484  
 A>Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-113 <KUR>  
 A:Cross-references: GB:BA000020; PIDN:BA078141.1; PID:917135595; GSPDB:GN00180  
 A:Experimental source: strain PCC 7120  
 C:Genetics:  
 A:Gene: alr7057  
 A:Genome: plasmid

Query Match 50.7% Score 37; DB 2; Length 113;  
 Best Local Similarity 60.0%; Pred. No. 14;  
 Matches 6; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AWYXXRGIRP 10  
 ||| |||  
 Db 64 AWYAKAGIEP 73

Search completed: September 13, 2002, 09:24:01  
 Job time: 776 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using SW model

Run on: September 13, 2002, 09:30:46 ; Search time 80.21 Seconds  
(without alignments)  
7.241 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_7\_21

Perfect score: 73

Sequence: 1 AMYXXRGIRPVGRFX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: SWISSPROT\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	70	95.9	87 1	PRRP_HUMAN
2	69	94.5	83 1	PRRP_RAT
3	69	94.5	98 1	PRRP_BOVIN
4	43	58.9	719 1	NRPL_YEAST
5	39	53.4	149 1	ENRN_BPT7
6	39	53.4	485 1	SYE_BACHD
7	38	53.4	676 1	EXLI_HUMAN
8	38	52.1	637 1	MTHA_DROME
9	38	52.1	2832 1	NDVB_RHIME
10	37	50.7	788 1	REC2_HABIN
11	36	49.3	402 1	EX7L_STRCO
12	35	47.9	179 1	RK6_GUTRH
13	35	47.9	383 1	CYCR_CHRVI
14	35	47.9	391 1	THIL_ZOORA
15	35	47.9	433 1	TBB2_GEOCN
16	35	47.9	460 1	SR54_HALNI
17	35	47.9	596 1	SDP_EIMBO
18	35	47.9	618 1	CIRL_CIFER
19	35	47.9	922 1	TYK2_HUMAN
20	35	47.9	1187 1	TYK2_HUMAN
21	34	46.6	224 1	YGC1_ECOLI
22	34	46.6	254 1	EMD_HUMAN
23	34	46.6	306 1	RS1_SYN6
24	34	46.6	328 1	RS1A_SYNY3
25	34	46.6	424 1	SAMB_SALTJ
26	34	46.6	476 1	YAAI_ECOLI
27	34	46.6	801 1	TFR2_HUMAN
28	34	46.6	943 1	SYL_PSEFL
29	34	46.6	987 1	EPB4_MOUSE
30	34	46.6	1046 1	HHUA_HABIN
31	34	46.6	1077 1	HGPA_HABIN
32	33	45.9	323 1	RBSC_HABIN
33	33	45.2	270 1	CH14_PHAVU

34	33	45.2	424 1	PIP_AERSO	P46547 aeromonas s
35	33	45.2	443 1	ARGA_ECOLI	P08205 escherichia
36	33	45.2	477 1	PROX_HUMAN	P50336 homo sapien
37	33	45.2	477 1	PROX_MOUSE	P51175 mus musculu
38	33	45.2	501 1	TDT_ONCHY	Q92089 oncorhynch
39	33	45.2	582 1	KICH_YEAST	P20485 saccharomyc
40	33	45.2	636 1	ENV_MCFP	P15073 mink cell f
41	33	45.2	737 1	CATA_CAUCR	Q31066 caulobacter
42	33	45.2	781 1	PBPB_HABIN	P45345 h penicillii
43	33	45.2	850 1	PBPA_ECOLI	P02918 escherichia
44	33	45.2	972 1	HGBA_HABDU	Q47952 haemophilus
45	33	45.2	972 1	HGBB_HABDU	Q47957 haemophilus

## ALIGNMENTS

```

RESULT 1
ID PRRP_HUMAN STANDARD; PRT; 87 AA.
AC P81277;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PRP31; Prolactin-
DE releasing peptide PRP20].
GN PRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; Pubmed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kiyada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RA "A prolactin-releasing peptide in the brain.";
RT Nature 393:272-276(1998).
RL [2]
RN TISSUE SPECIFICITY.
RP Pubmed=10498338;
RX Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Hinuma S.,
RA Sekiguchi M., Kiyada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RA "Tissue distribution of prolactin-releasing peptide (PRP) and its
RA receptor.";
RT Regul. Pept. 83:1-10(1999).
CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL: AB015419; BAA29027.1; -.
CC DR MIM: 602663; -.
CC KW Hormone; Amidation; Signal.
CC FT SIGNAL 1 22
CC FT PEPTIDE 23 53 BY SIMILARITY.
CC FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRP31.
CC FT MOD_RES 53 53 PROLACTIN-RELEASING PEPTIDE PRP20.
CC FT MOD_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).
CC SQ SEQUENCE 87 AA: 9639 MW: 229423550CF981B CRC64;

```

Query Match 95.9%; Score 70; DB 1; Length 87;  
Best Local Similarity 85.7%; Pred. No. 8.5e-07;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXGIRPVGRF 14  
DB 40 AMYASGIRPVGRF 53

## RESULT 2

PRRP\_RAT STANDARD: PRT: 83 AA.

AC P81278:  
30-MAY-2000 (Rel. 39, Created)  
30-MAY-2000 (Rel. 39, Last sequence update)  
01-MAR-2002 (Rel. 41, Last annotation update)  
DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].  
DE releasing peptide PrRP20].  
GN PRH.

OS Rattus norvegicus (Rat).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Rattus. NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.

RC TISSUE=Brain;  
RX MEDLINE=98268781; PubMed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RT "A proactin-releasing peptide in the brain.";  
RL Nature 393:272-276(1998).  
RN [2]  
RP TISSUE SPECIFICITY.

RX PubMed=10498338;  
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
RT "Tissue distribution of proactin-releasing peptide (PrRP) and its receptor.";  
RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: AB015418; BAA29026.1;  
KW Hormone; Amidation; Signal;  
FT SIGNAL 1 21  
FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.  
FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.  
FT MOD\_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 83 AA; 9215 MW; DDC75A264EEB4F29 CRC64;

Query Match 94.5%; Score 69; DB 1; Length 83;  
Best Local Similarity 85.7%; Pred. No. 1.3e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXGIRPVGRF 14  
DB 39 AMYTGIRPVGRF 52

## RESULT 3

PRRP\_BOVIN STANDARD: PRT: 98 AA.

AC P81264:  
30-MAY-2000 (Rel. 39, Created)  
30-MAY-2000 (Rel. 39, Last sequence update)  
01-MAR-2002 (Rel. 41, Last annotation update)  
DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].  
DE releasing peptide PrRP20].  
GN PRH.

OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos. NCBI\_TaxID=9913;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RC TISSUE=Brain;  
RX MEDLINE=98268781; PubMed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RT "A proactin-releasing peptide in the brain.";  
RL Nature 393:272-276(1998).  
CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
CC EMBL: AB015417; BAA29025.1;  
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
FT SIGNAL 1 22  
FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 94.5%; Score 69; DB 1; Length 98;  
Best Local Similarity 85.7%; Pred. No. 1.5e-06;  
Matches 12; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 AMYXXGIRPVGRF 14  
DB 40 AMYAGGIRPVGRF 53

## RESULT 4

NRPL\_YEAST STANDARD: PRT: 719 AA.

AC P32770; Q12228;  
01-OCT-1993 (Rel. 27, Created)  
01-NOV-1997 (Rel. 35, Last sequence update)  
01-MAR-2002 (Rel. 41, Last annotation update)  
DE Asparagine-rich protein (ARP protein).

OS Saccharomyces cerevisiae (baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes; Saccharomycetales; Saccharomycetaceae; Saccharomyces. NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-AH22;

```

RX MEDLINE-93247548; PubMed-8483449;
RA Weiner E.P., Rao E., Brendel M.;
RT "Molecular structure and genetic regulation of SFA, a gene
RT responsible for resistance to formaldehyde in Saccharomyces
RT cerevisiae, and characterization of its protein product.";
RL Mol. Gen. Genet. 237:351-358(1993).
RN [2].
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RA POL1 T.M.;
RL Submitted (NOV-1995) to the EMBL/Genbank/DBD databases.
CC -1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68020; CAA48159.1; -
DR EMBL: Z67750; CAA91579.1; -
DR EMBL: Z74215; CAA98741.1; -
DR PIR: S31139; S31139.
DR HSSP: P04170; 6RXN.
DR SCG: S0002326; NRPL.
DR InterPro: IPR001876; Znf-RanBP.
DR Pfam: PF00076; Fm; 1.
DR Pfam: PF00641; Zf-RanBP; 2.
DR SMART: SM00360; RRM; 1.
DR SMART: SM00547; Znf_RBZ; 2.
DR PROSITE: PS00103; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PS01358; ZF-RANBP2_1; 2.
DR PROSITE: PS50199; ZF-RANBP2_2; 2.
KW Nuclear protein; zinc-finger; RNA-binding; Repeat.
FT DOMAIN 226 322 RNA-BINDING (RRM).
FT 2N-FING 355 384 RANBP2-TYPE 1.
FT 2N-FING 581 610 RANBP2-TYPE 2.
FT DOMAIN 490 564 ASN-RICH.
FT CONFLICT 493 493 I -> N (IN REF. 1).
SQ SEQUENCE 719 AA; 79299 MW; ADA9BC09F582669 CRC64;

Query Match 58.9%; Score 43; DB 1; Length 719;
Best Local Similarity 50.0%; Pred. NO. 1.3;
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPYG 12
   :|: |::|||
Db 244 SMFTQYGVRPVG 255

RESULT 5
ENRN_BPT7 STANDARD; PRT; 149 AA.
AC P00641;
DT 21-JUL-1986 (Rel. 01, Created)
DT 21-JUL-1986 (Rel. 01, Last sequence update)
DT 01-NOV-1995 (Rel. 32, Last annotation update)
DE Endonoxylribonuclease I (EC 3.1.21.2) (Endonuclease).
GN 3.
OS Bacteriophage T7,
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;
OC T7-like phages.
OX NCBI_TaxID=10760;
OX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE-83241725; PubMed-6864790;
RA Dunn J.J., Studier F.W.;
```

```

RT "Complete nucleotide sequence of bacteriophage T7 DNA and the
RT locations of T7 genetic elements.";
RL J. Mol. Biol. 166:477-535(1983).
RN [2].
RP SEQUENCE FROM N.A.
RX MEDLINE-82078034; PubMed-7310871;
RA Dunn J.J., Studier F.W.;
RT "Nucleotide sequence from the genetic left end of bacteriophage T7
RT DNA to the beginning of gene 4.";
RL J. Mol. Biol. 148:303-330(1981).
CC -1- FUNCTION: ENDONOXYLRIPONUCLEASE I, WHICH IS EXPRESSED IN THE LATE
CC STAGE, IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN
CC OF HOST DNA. IN THE EARLY STAGE OF INFECTION, T7 DNA REPLICATES AS
CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA
CC LINEAR CONCATAMERS SEVERAL GENOMES IN LENGTH. THE GENOME 3 PRODUCT
CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE CONCATAMERS.
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-
CC phosphooligonucleotide end-products.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: V01127; CAA24345.1; -
DR EMBL: V01146; CAA24402.1; -
DR PIR: A00785; NEBP37.
DR PIR: S42301; S42301.
KW Hydrolyase; Nuclease; Endonuclease.
SQ SEQUENCE 149 AA; 17172 MW; D092AA28E3743BC1 CRC64;

Query Match 53.4%; Score 39; DB 1; Length 149;
Best Local Similarity 57.1%; Pred. NO. 1.6;
Matches 8; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1 AMYXXRGIRPYG 14
   | |::||| | |
Db 2 AGYGAKKIRKYGAF 15

RESULT 6
SYE_BACHD STANDARD; PRT; 485 AA.
ID SYE_BACHD
AC Q9KGF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamy1-cRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)
DE (GluRS).
GN GLTX OR BH0109.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group;
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
OX [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;
RX MEDLINE-20512582; PubMed-11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kihara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC diphosphate + L-glutamyl-tRNA(Glu).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
```

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP001507; BAB03828.1; -  
 DR InterPro: IPR000924; tRNA-synt\_1c.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR Pfam: PF00749; tRNA-synt\_1c.1.  
 DR PRINTS: PR00987; TRNASYNTHLU.  
 DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I.1.  
 DR Aminoacyl-tRNA synthetase: Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 11 21 "HIGH" REGION.  
 FT SITE 252 256 "KMSKS" REGION.  
 FT BINDING 255 255 ATP (BY SIMILARITY).  
 FT SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;  
 SQ  
 Query Match 53.4%; Score 39; DB 1; Length 485;  
 Best Local Similarity 58.3%; Pred. No. 5.2;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 Oy 3 YXXRGIRPYGRF 14  
 Db 146 YKAKGIRPYGRF 157  
 RESULT 7  
 EXTL\_HUMAN  
 ID EXTL\_HUMAN STANDARD; PRT; 676 AA.  
 AC Q92935;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Exostosin-like 1 (Exostosin-L) (Multiple exostosin-like protein).  
 GN EXTL1 OR EXTL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=97189339; PubMed=9037597;  
 RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
 RT "Identification and localization of the gene for EXTL1, a third member  
 RT of the multiple exostosin gene family.";  
 RL Genome Res. 7:10-16(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Xu L., Deng H.X., Xia J.H., Pan O., Liu C.Y.;  
 RT "Mutations of the EXT genes in hereditary multiple exostosin in  
 RT Chinese";  
 RL Submitted (Aug-1998) to the EMBL/GenBank/DBD databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Muijs W., Spieker N., Van Roy N., De Paeppe A., De Boule K.,  
 RA Millems P.J., Van Hul W., Versteeg R., Speleman F.;  
 RT "Refined physical mapping and genomic structure of the EXTL1 gene";  
 RL Submitted (May-1999) to the EMBL/GenBank/DBD databases.  
 CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way

-----  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67191; AAC51141.1; -  
 DR EMBL: AF083633; AAD02840.1; -  
 DR EMBL: AF083623; AAD02840.1; JOINED.  
 DR EMBL: AF083624; AAD02840.1; JOINED.  
 DR EMBL: AF083625; AAD02840.1; JOINED.  
 DR EMBL: AF083626; AAD02840.1; JOINED.  
 DR EMBL: AF083627; AAD02840.1; JOINED.  
 DR EMBL: AF083628; AAD02840.1; JOINED.  
 DR EMBL: AF083629; AAD02840.1; JOINED.  
 DR EMBL: AF083630; AAD02840.1; JOINED.  
 DR EMBL: AF083631; AAD02840.1; JOINED.  
 DR EMBL: AF083632; AAD02840.1; JOINED.  
 DR EMBL: AF153980; AAF73172.1; -  
 DR EMBL: AF153991; AAF73172.1; JOINED.  
 DR MIM: 601738; -  
 DR InterPro: IPR004263; Exostosin.  
 DR Pfam: PF03016; Exostosin.1.  
 KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.  
 FT TRANSMEM 10 30 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT CARBOHYD 269 269 N-LINKED (GLCNAC... ) (POTENTIAL).  
 FT ANTI-ONCOGENE 10 30 (POTENTIAL).  
 FT SEQUENCE 676 AA; 74673 MW; B5E006A8762E3633 CRC64;  
 SQ  
 Query Match 53.4%; Score 39; DB 1; Length 676;  
 Best Local Similarity 53.8%; Pred. No. 7.3;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 Oy 2 WYXXRGIRPYGRF 14  
 Db 407 YVYQGSRRPGRF 419  
 RESULT 8  
 MTHA\_DROME  
 ID MTHA\_DROME STANDARD; PRT; 637 AA.  
 AC Q9W0R5;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Probable G-protein-coupled receptor Mth-like 10 precursor (Methuselah-  
 DE like 10 protein).  
 GN MTHL10 OR CG17061.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxId=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-Berkeley;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanalides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Chang M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Arrill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Bailew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borrova D., Botchan M.R., Bouck J., Brockstein P., Brottier P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,  
 RA Flosser C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,



FT	CARBOHYD	142	142	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	152	152	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	157	157	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	198	198	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
FT	CARBOHYD	223	223	N-LINKED (GLCNAC. . .)	(POTENTIAL).				
SO	SEQUENCE	637 AA;	72988 MM;	CC5FA399B15CE777	CRC64;				
Query Match		52.1%	Score 38;	DB 1;	Length 637;				
Best Local Similarity		46.2%	Pred. No. 11;						
Matches		6;	Conservative	2;	Mismatches				
				5;	Indels				
				0;	Gaps				
				0;					
OY	2 WYXXRGIRPVGRF	14							
		1: 11 1: 11							
DB	323 WHNFRGTGRINRF	335							
RESULT									
9									
NDVB_RHIME	STANDARD;	PRT;	2832	AA.					
AC	P20471;								
DT	01-FEB-1991 (Rel. 17, Created)								
DT	01-MAR-2002 (Rel. 41, Last sequence update)								
DT	01-MAR-2002 (Rel. 41, Last annotation update)								
DE	Protein ndvb.								
NDVB	OR R03286 OR SMC04382.								
OS	Rhizobium meliloti (Sinorhizobium meliloti).								
OC	Bacteriae; Proteobacteria; alpha subdivision; Rhizobiaceae group;								
OC	Rhizobiaceae; Sinorhizobium.								
OX	NCBI_TaxID=382;								
[1]									
RP	SEQUENCE FROM N.A.								
RX	MEDLINE=90153914; PubMed=2154461;								
RA	Jepl L., Dlyan T., Dltta G.S., Helinski D.R., Stanfield S.W.;								
RT	"The ndvb locus of Rhizobium meliloti encodes a 319-kDa protein								
RT	involved in the production of beta-(1-->2)-glucan.";								
RL	J. Biol. Chem. 265:2843-2851(1990).								
[2]									
RN	SEQUENCE FROM N.A.								
RP	STRAIN=1021;								
RC	MEDLINE=21396507; PubMed=11481430;								
RX	Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,								
RA	Boisard F., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,								
RA	Godrie T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,								
RA	Polh T., Portetalle D., Puhler A., Purnelle B., Ramsperger U.,								
RA	Reard C., Thebault P., Vandenbol M., Weidner S., Gallibert F.;								
RT	"Analysis of the chromosome sequence of the legume symbiont								
RT	Sinorhizobium meliloti strain 1021.";								
RL	Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).								
CC	-1- FUNCTION: INVOLVED IN THE PRODUCTION OF BETA-(1,2)-GLUCAN.								
CC	IT IS INVOLVED NOT ONLY IN INVASION BUT ALSO IN BACTEROID								
CC	DEVELOPMENT.								
CC	-1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.								
CC	-1- SIMILARITY: TO A.TIMERACIENS CHVB.								
CC	-----								
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration								
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -								
CC	the European Bioinformatics Institute. There are no restrictions on its								
CC	use by non-profit institutions as long as its content is in no way								
CC	modified and this statement is not removed. Usage by and for commercial								
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>								
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).								

FT CONFLICT 1187 1187 L -> F (IN REF. 1).  
 FT CONFLICT 2602 2602 D -> G (IN REF. 1).  
 SQ SEQUENCE 2832 AA; 315746 MW; 171EA89F03A936F2 CRC64;

Query Match 52.1%; Score 38; DB 1; Length 2832;  
 Best Local Similarity 50.0%; Pred. No. 49;  
 Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 WYXXRGIRPV 11  
 DB 1088 WYTRGLEP 1097

RESULT 10  
 RECD\_HAEIN STANDARD; PRT; 788 AA.  
 AC P44408;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Recombination protein 2.  
 GN REC2 OR REC-2 OR HI0061.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellales;  
 OC Haemophilus.  
 OC NCBI\_Taxid=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BC200;  
 RX MEDLINE-94341577; PubMed-8063112;  
 RA Clifton S.W., McCarthy D., Roe B.A.;  
 RT \*Sequence of the rec-2 locus of Haemophilus influenzae: homologues to  
 RL Gene 146:95-100(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KW20 / ATCC 51907;  
 RX MEDLINE-95350630; PubMed-7542800;  
 RA Kleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Leisegang A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
 RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhmann J.L., Geoghegan N.S.M.,  
 RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT \*Whole-genome random sequencing and assembly of Haemophilus  
 RT Influenzae Rd.;  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: MIGHT CONTRIBUTE TO TRANSFORMATION AS A MEMBER OF A  
 CC MEMBRANE BOUND PORE COMPLEX AT THE BASE OF THE TRANSFORMASOME. IT  
 CC COULD DIRECTLY INTERACT WITH TRANSFORMING DNA DURING TRANSLLOCATION  
 CC INDIRECTLY BY PARTICIPATING IN THE ASSEMBLY OF THE PORE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 CC (Potential).  
 CC -1- SIMILARITY: TO B.SUBTILIS COMEC, N.CONORRHORAE COMA, AND E.COLI  
 CC YCAL.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; L20805; AAC13733.1; -  
 DR EMBL; U32691; AAC21739.1; -  
 DR TIGR; H10061; -  
 DR InterPro: IPR001279; Beta.Lactam\_mec.  
 DR Pfam: PF00753; Lactamase\_B; 1.

KW Transport; Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 3 23 POTENTIAL.  
 FT TRANSMEM 27 47 POTENTIAL.  
 FT TRANSMEM 51 71 POTENTIAL.  
 FT TRANSMEM 226 246 POTENTIAL.  
 FT TRANSMEM 263 283 POTENTIAL.  
 FT TRANSMEM 313 333 POTENTIAL.  
 FT TRANSMEM 367 387 POTENTIAL.  
 FT TRANSMEM 401 421 POTENTIAL.  
 FT TRANSMEM 434 454 POTENTIAL.  
 FT TRANSMEM 456 476 POTENTIAL.  
 FT TRANSMEM 508 528 POTENTIAL.  
 FT CONFLICT 748 788  
 FT  
 FT  
 SQ SEQUENCE 788 AA; 89355 MW; F31104595CBAE7A CRC64;  
 IKIGTMRATFIR (IN REF. 1).  
 -> GRKRCFGASAGFEFSRPIRNPASSHKIPPLVCACNMI

Query Match 50.7%; Score 37; DB 1; Length 788;  
 Best Local Similarity 54.5%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 WYXXRGIRPV 12  
 DB 156 WYFSKGTAVG 166

RESULT 11  
 EX7L\_STRCO STANDARD; PRT; 402 AA.  
 ID EX7L\_STRCO  
 AC Q9FBN3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
 DE (Exonuclease VII large subunit).  
 GN XSEA OR SCK7.29C.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomycetes.  
 OC NCBI\_Taxid=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D., Cerdano A.M., Parkhill J., Barrell B.G.,  
 RA Rajandream M.A.;  
 RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
 CC -1- FUNCTION: BIDIRECTIONALLY DEGRADATES SINGLE-STRANDED DNA INTO LARGE  
 CC ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
 CC INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5'- to 3'-  
 CC or 3'- to 5'-direction to yield 5'-phosphomononucleotides.  
 CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
 CC SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AL391754; CAC05901.1; -  
 DR InterPro: IPR003753; Exonuc\_VII\_L.  
 DR InterPro: IPR002309; tRNA-syn\_2.  
 DR Pfam; PF02601; Exonuc\_VII\_L; 1.  
 DR Pfam; PF01336; tRNA\_ant1; 1.  
 DR Hydroxase; Nuclease; Exonuclease.  
 SQ SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 49.3%; Score 36; DB 1; Length 402;  
Best Local Similarity 42.1%; Pred. No. 17;  
Matches 8; Conservative 1; Mismatches 2; Indels 8; Gaps 1;  
OY 2 WYXXRG-----IRPVG 12  
|||  
Db 91 WYAPRGQLSLRAAEIKPVG 109

RESULT 12  
RK6\_GUITH  
ID RK6\_GUITH STANDARD; PRT: 179 AA.  
AC 046908;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Chloroplast 50S ribosomal protein L6.  
GN RPL6.  
OS Guillardia theta (Cryptomonas ph1).  
OC Chloroplast.  
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
OX NCBI\_TaxID=55529;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97283757; PubMed=9137835;  
RA Wang S.L., Liu X.-Q., Douglas S.E.;  
RT "The large ribosomal protein gene cluster of a cryptomonad plastid:  
RT gene organization, sequence and evolutionary implications.";  
RL Blochem. Mol. Biol. Int. 41:1035-1044(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99128221; PubMed=9929392;  
RA Douglas S.E., Penny S.L.;  
RT "The plastid genome of the cryptophyte alga, Guillardia theta:  
RT complete sequence and conserved synteny groups confirm its common  
RT ancestry with red algae.";  
RL J. Mol. Evol. 46:236-244(1999).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
CC LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE  
CC PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AF041468; AAC35717.1; -;  
DR HSSP; P02391; RL6.  
DR InterPro: IPR000702; Ribosomal\_L6.  
DR InterPro: IPR003358; Ribosomal\_L6\_1.  
DR Pfam; PF00347; Ribosomal\_L6; 1.  
DR PRINTS; PR00059; RIBOSOMAL\_L6.  
DR PRODOM; PD002236; RIBOSOMAL\_L6; 1.  
DR PROSITE; PS00525; RIBOSOMAL\_L6\_1; 1.  
KW Ribosomal protein; Chloroplast.  
SQ SEQUENCE 179 AA; 19527 MW; 8B4COBDB015ZAD24 CRC64;

Query Match 47.9%; Score 35; DB 1; Length 179;  
Best Local Similarity 58.3%; Pred. No. 11;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
OY 3 YXXRGIRPVGRF 14  
|:|||||  
Db 157 YXGKGRVGEF 168

RESULT 13

CYCR\_CHRYI  
ID CYCR\_CHRYI STANDARD; PRT: 363 AA.  
AC 082947;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 30-MAY-2000 (Rel. 39, Last annotation update)  
DE Photosynthetic reaction center cytochrome C subunit precursor.  
GN PufC.  
OS Chromatium vinosum.  
OC Bacteria; Proteobacteria; gamma subdivision; Chromatiaceae;  
OC Allochromatium.  
OX NCBI\_TaxID=1049;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=D;  
RA Corson G.E., Nagashima K.V., Matsura K., Sakuragi Y., Ruwanti W.,  
RA Qiu H., Allen R., Knaff D.B.;  
RT "Primary structure of genes encoding light-harvesting and reaction  
RT center proteins from Chromatium vinosum.";  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: THE REACTION CENTER OF PURPLE BACTERIA CONTAIN A  
CC TIGHTLY BOUND CYTOCHROME MOLECULE WHICH RE-REDUCES THE PHOTO  
CC OXIDIZED PRIMARY ELECTRON DONOR.  
CC -1- SUBCELLULAR LOCATION: ATTACHED TO THE MEMBRANE BY A LIPID  
CC ANCHOR (BY SIMILARITY).  
CC -1- PTM: BINDS FOUR HEME GROUPS PER MOLECULE.  
CC -1- SIMILARITY: HIGH, WITH OTHER PHOTOSYNTHETIC REACTION CENTER  
CC CYTOCHROME C SUBUNITS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL; AB011811; BAA32742.1; -;  
DR HSSP; P07173; 6PRC.  
DR InterPro: IPR003158; CytoC\_RC.  
DR InterPro: IPR000345; CytoC\_heme\_bind.  
DR Pfam; PF02276; CytoC\_RC; 1.  
DR PROSITE; PS00190; CYTOCHROME\_C; 4.  
KW Electron transport; Photosynthesis; Reaction center; Heme;  
KW Membrane; Lipoprotein; Duplication; Signal.  
FT SIGNAL 1 22  
FT CHAIN 23 383  
FT  
FT LIPID 23 23  
FT BINDING 107 107  
FT BINDING 110 110  
FT METAL 111 111  
FT  
FT BINDING 152 152  
FT BINDING 155 155  
FT METAL 156 156  
FT  
FT BINDING 247 247  
FT BINDING 250 250  
FT METAL 251 251  
FT  
FT BINDING 307 307  
FT BINDING 310 310  
FT METAL 311 311  
SQ SEQUENCE 383 AA; 41522 MW; 96BCD91FF1B9AE7E CRC64;

Query Match 47.9%; Score 35; DB 1; Length 383;  
Best Local Similarity 38.5%; Pred. No. 25;  
Matches 5; Conservative 3; Mismatches 5; Indels 0; Gaps 0;  
OY 1 AMYXXRGIRPVGR 13

DB 269 AWAIHRVDING 281

```

|||||
RESULT 14
THL_ZOORA STANDARD; PRT; 391 AA.
ID THL_ZOORA
AC P07097;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 01-OCT-1996 (Rel. 34, Last annotation update)
DE Acetyl-CoA acetyltransferase (EC 2.3.1.9) (acetoacetyl-CoA thiolase).
GN PHB.
OS Zoogloea ramigera.
OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;
OC Zoogloea.
OX NCBI_TaxID=350;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 19623 / I-16-M;
RX MEDLINE=87083504; PubMed=2878929;
RA Peoples O.P., Masumne S., Walsh C.T., Sinskey A.J.;
RT "Biosynthetic thiolase from Zoogloea ramigera. Evidence for a
RT mechanism involving Cys-378 as the active site base.";
RL J. Biol. Chem. 266:8369-8375(1991).
CC -1- CATALYTIC ACTIVITY: 2 acetyl-CoA + acetoacetyl-CoA.
CC -1- PATHWAY: FIRST STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE THIOLEASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: J02631; AAA27706.1; ALT_SEQ.
DR PIR: A26121; XXGZAC.
DR PIR: A2754; A27754.
DR HSSP: P27796; IAFY.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF02803; thiolase; 1.
DR PROSITE: PS00098; THIOLEASE_1; 1.
DR PROSITE: PS00099; THIOLEASE_3; 1.
DR PROSITE: PS00737; THIOLEASE_2; 1.
KW Transferrase; Acyltransferase; PHB biosynthesis.
FT INIT_MEP 0
FT ACT_SITE 88 88 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 377 377 BASE.
FT MUTAGEN C>G: LOSS OF ACTIVITY.
SQ SEQUENCE 391 AA; 40342 MW; 6D2351A1BC0E4EDD CRC64;

```

Query Match 47.9%; Score 35; DB 1; Length 391;  
 Best Local Similarity 75.0%; Pred. No. 25;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 6 RGIPIYGR 13  
 DB 266 RGIPIYGR 273

```

|||||
RESULT 15
TBB2_GEOCN STANDARD; PRT; 453 AA.
ID TBB2_GEOCN
AC P32925;
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta-2 chain.
GN Geotrichum candidum (Oospora lactis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Galactomycetes.
OX NCBI_TaxID=27317;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92079883; PubMed=1836049;
RA Gold S.E., Casale W.L., Keen N.T.;
RT "Characterization of two beta-tubulin genes from Geotrichum
RT candidum.";
RL Mol. Gen. Genet. 230:104-112(1991).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch)
CC -----
DR EMBL: S69627; AAB20557.1; ALT_SEQ.
DR PIR: S18597; S18597.
DR InterPro: IPR002453; Beta_tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_FtsZ.
DR Pfam: PF0161; TUBULIN.
DR PRINTS: PR0161; TUBULIN.
DR PROSITE: PS00227; TUBULIN_1.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG; 1.
KW Microtubules; GTP-binding; Multigene family.
FT NP_BIND 142 148 GTP (POTENTIAL).
SQ SEQUENCE 453 AA; 50399 MW; 2EA9D8A0246E0371 CRC64;

```

Query Match 47.9%; Score 35; DB 1; Length 453;  
 Best Local Similarity 50.0%; Pred. No. 25;  
 Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 WYXXRGIRPV 11  
 DB 399 WYXGMEPV 408

Search completed: September 13, 2002, 09:30:47  
 Job time: 1137 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:20 ; Search time 311.85 Seconds  
(without alignments)  
8.321 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_7\_21

Perfect score: 73

Sequence: 1 AMYXXRGIRPVGRFX 15

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

SPREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mnc:\*  
8: sp\_organelle:\*  
9: sp\_phase:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriaph:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	84.9	117	13	09W624
2	46.5	63.7	420	2	093LZ7
3	40	54.8	153	9	09P133
4	40	54.8	184	2	086838
5	40	54.8	540	10	091G20
6	39	53.4	179	2	09K307
7	39	53.4	256	16	092VE9
8	39	53.4	425	16	0986U6
9	39	53.4	555	3	000050
10	38	52.1	194	16	09A9U0
11	38	52.1	238	17	09YEH0
12	38	52.1	333	16	09PH76
13	38	52.1	398	16	09RR92
14	38	52.1	481	5	095WT7
15	38	52.1	536	5	095YV4
16	38	52.1	536	5	095PK7

17	38	52.1	536	5	095PK6	095pk6 drosophila
18	38	52.1	536	5	095PK5	095pk5 drosophila
19	38	52.1	536	5	095NR7	095nr7 drosophila
20	38	52.1	536	5	095NR6	095nr6 drosophila
21	38	52.1	536	5	095NR7	095nr7 drosophila
22	38	52.1	536	5	095NR0	095nr0 drosophila
23	37	50.7	114	11	063104	063104 ratius norv
24	37	50.7	160	10	09LMD3	09lmd3 oryza sativ
25	37	50.7	234	11	063112	063112 ratius norv
26	37	50.7	293	16	09PHR2	09phr2 campylobact
27	37	50.7	300	16	092SA6	092sa6 rhizobium m
28	37	50.7	302	16	098BC2	098bc2 rhizobium l
29	37	50.7	320	16	091142	091142 pseudomonas
30	37	50.7	332	5	09W043	09w043 pseudomonas
31	37	50.7	338	16	092X86	092x86 rhizobium m
32	37	50.7	343	16	092UH4	092uh4 rhizobium m
33	37	50.7	348	5	09W1G4	09w1g4 drosophila
34	37	50.7	357	5	095U70	095u70 drosophila
35	37	50.7	395	2	054095	054095 streptomyc
36	37	50.7	769	2	068843	068843 streptomyc
37	37	50.7	885	16	092NX4	092nx4 rhizobium m
38	37	50.7	1450	11	054728	054728 ratius norv
39	36.5	50.0	472	16	0910V9	0910v9 pseudomonas
40	36.5	50.0	541	2	046610	046610 desulfovibr
41	36	49.3	163	2	045285	045285 bacillus l1
42	36	49.3	220	16	091022	091022 pseudomonas
43	36	49.3	276	16	053929	053929 mycobacteri
44	36	49.3	293	2	053855	053855 splioplasma
45	36	49.3	302	16	P96274	P96274 mycobacteri

## ALIGNMENTS

RESULT	ID	Query Match	Best Local Similarity	Matches	Score	DB	Length
1	09W624	PRELIMINARY;	69.28;	9;	62;	13;	117;
09W624	09W624	PRELIMINARY;	69.28;	9;	62;	13;	117;
AC	09W624	PRELIMINARY;	69.28;	9;	62;	13;	117;
DT	01-NOV-1999	(TREMBLrel. 12, Created)	69.28;	9;	62;	13;	117;
DT	01-NOV-1999	(TREMBLrel. 12, Last sequence update)	69.28;	9;	62;	13;	117;
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	69.28;	9;	62;	13;	117;
DE	C-RF AMIDE.		69.28;	9;	62;	13;	117;
OS	Carassius auratus (Goldfish).		69.28;	9;	62;	13;	117;
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		69.28;	9;	62;	13;	117;
CC	Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;		69.28;	9;	62;	13;	117;
OC	Cypriniformes; Cyprinidae; Carassius.		69.28;	9;	62;	13;	117;
OX	NCBI_TaxID=7957;		69.28;	9;	62;	13;	117;
RN	[1]		69.28;	9;	62;	13;	117;
RP	SEQUENCE FROM N.A.		69.28;	9;	62;	13;	117;
RC	TISSUE=BRIN;		69.28;	9;	62;	13;	117;
RA	Satoka H., Minakata H., Fujimoto M.;		69.28;	9;	62;	13;	117;
RT	"Carassius auratus (C-RF amide)." ;		69.28;	9;	62;	13;	117;
RL	Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.		69.28;	9;	62;	13;	117;
DR	EMBL; AB020024; BAA6662.1;		69.28;	9;	62;	13;	117;
SO	SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;		69.28;	9;	62;	13;	117;
QY	2 WYXXRGIRPVGRF 14		69.28;	9;	62;	13;	117;
DB	63 WYXGVRPVGRF 75		69.28;	9;	62;	13;	117;
RESULT	2		69.28;	9;	62;	13;	117;
093LZ7	093LZ7	PRELIMINARY;	69.28;	9;	62;	13;	117;
AC	093LZ7	PRELIMINARY;	69.28;	9;	62;	13;	117;
DT	01-DEC-2001	(TREMBLrel. 19, Created)	69.28;	9;	62;	13;	117;
DT	01-DEC-2001	(TREMBLrel. 19, Last sequence update)	69.28;	9;	62;	13;	117;
DT	01-DEC-2001	(TREMBLrel. 19, Last annotation update)	69.28;	9;	62;	13;	117;

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
 GN AUR2B.  
 OS Streptomyces aureofaciens.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1894;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCM3239;  
 RA Kormaneec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
 RT "Cloning and characterization of a new polypeptide gene cluster in  
 RT Streptomyces aureofaciens CCM3239."  
 RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF033994; AAK61719.1; -  
 SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 63.7%; Score 46.5; DB 2; Length 420;  
 Best Local Similarity 52.6%; Pred. No. 1.5;  
 Matches 10; Conservative 1; Mismatches 3; Indels 5; Gaps 1;

OY 1 AMYXX-----RGIRPVGRF 14  
 ||:|||||  
 Db 37 AMMAAVLRGSGIRPVGRF 55

RESULT 3  
 ID 09T133 PRELIMINARY; PRT; 153 AA.  
 AC 09T133;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE ENDONUCLEASE.  
 DE GN 3.  
 OS Bacteriophage phlyeo3-12.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
 OC T7-like phages.  
 OX NCBI\_TaxID=110457;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Pajunen M.I.;  
 RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=21125594; PubMed=11222590;  
 RA Pajunen M.I., Kiljunen S.J., Soederholm M.E.L., Skurnik M.;  
 RT "Complete genomic sequence of the lytic bacteriophage phlyeo3-12 of  
 RT Yersinia enterocolitica serotype O:3.";  
 RL J. Bacteriol. 183:1928-1937(2001).  
 DR EMBL: AJ251805; CAB63604.1; -  
 SQ SEQUENCE 153 AA; 17640 MW; 211571BBDE6C641D CRC64;

Query Match 54.8%; Score 40; DB 9; Length 153;  
 Best Local Similarity 58.3%; Pred. No. 8;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 3 YXXRGIRPVGRF 14  
 ||:|||||  
 Db 5 YAAAGYRKVGAF 16

RESULT 4  
 ID 086838 PRELIMINARY; PRT; 184 AA.  
 AC 086838;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE PUTATIVE MEMBRANE PROTEIN.  
 GN SC9A10.05C.  
 OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Harris D., Taylor K.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RA Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,  
 RA Knasahl H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL031260; CAA20292.1; -  
 SQ SEQUENCE 184 AA; 20178 MW; 58806A19F8FD596 CRC64;

Query Match 54.8%; Score 40; DB 2; Length 184;  
 Best Local Similarity 50.0%; Pred. No. 9.8;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGRF 14  
 ||:|||||  
 Db 90 AMVSAHGVRQVGPF 103

RESULT 5  
 ID 09LG20 PRELIMINARY; PRT; 540 AA.  
 AC 09LG20;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE EST AU070209(R3722) CORRESPONDS TO A REGION OF THE PREDICTED  
 DE GENE.  
 DE OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPPONBARE;  
 RA Sasaki T., Matsumoto T., Yamamoto K.;  
 RT "Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC  
 RT clone:PO702P03.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: CATALYZES THE COVALENT ATTACHMENT OF UBIQUITIN TO OTHER  
 CC PROTEINS (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: ATP + UBIQUITIN + PROTEIN LYSINE - AMP +  
 CC PYROPHOSPHATE + PROTEIN N-UBIQUITYLlysine.  
 CC -1- PATHWAY: SECOND STEP IN UBIQUITIN CONJUGATION.  
 CC -1- MISCELLANEOUS: A CYSTEINE RESIDUE IS REQUIRED FOR UBIQUITIN-  
 CC THIOLESTER FORMATION (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE UBIQUITIN-CONJUGATING ENZYME FAMILY.  
 DR EMBL: AP002481; BAA96583.1; -  
 DR HSSP: P06104; IAYZ.  
 DR InterPro: IPR000608; UBQ\_conjugat.  
 DR Pfam: PF00179; UBQ\_con.1.  
 DR SMART: SM00212; UBCC; 1.  
 DR PROSITE: PS50127; UBIQUITIN\_CONJUGAT\_2; 1.  
 KW ligase; Ubiquitin conjugation.  
 SQ SEQUENCE 540 AA; 60487 MW; 5DE1FAEB75A86E CRC64;

Query Match 54.8%; Score 40; DB 10; Length 540;  
Best Local Similarity 60.0%; Pred. No. 32;  
Matches 6; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 AMYXXRGIRP 10  
DB 73 AMYRGIRLP 82

## RESULT 6

09K307 PRELIMINARY; PRT; 179 AA.  
AC 09K307;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
DT 01-JUN-2001 (TREMblrel. 17, Last annotation update)  
DE PUTATIVE ACETYLTRANSFERASE.  
GN 25CG4.07C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Seeger K.J., Harris D.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RX MEDLINE=97000351; PubMed=8843436;  
RA Redenbach M., Kieser H.M., Denapate D., Eichner A., Cullum J.,  
RA Knaeuhl H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL360034; CAB95978.1; -  
DR InterPro; IPR000182; Acetyltransf\_GCN5.  
DR Pfam; PF00583; Acetyltransf; 1.  
KW Transferase.  
SQ SEQUENCE 179 AA; 19784 MW; 0693898A177C64C8 CRC64;

Query Match 53.4%; Score 39; DB 2; Length 179;  
Best Local Similarity 58.3%; Pred. No. 15;  
Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

QY 1 AMYXXRGIRP 12  
DB 138 AMYRGIRRTG 149

## RESULT 7

092VE9 PRELIMINARY; PRT; 256 AA.  
AC 092VE9;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE HYPOTHEICAL PROTEIN SMB21253.  
GN SMB21253.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Plasmid pSymb (megaplasmid 2).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]

RP SEQUENCE FROM N.A.  
RC STRAIN-1021;  
RX MEDLINE=21396508; PubMed=11481431;  
RA Fian T.M., Weidner S., Wong K., Buhrmester J., Chai P.,  
RA Vorhoefer F.J., Hernandez-Lucas I., Becker A., Cowie A., Guzy J.,  
RA Golding B., Fuenher A.;  
RT "The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-  
fixing endosymbiont Sinorhizobium meliloti.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
DR EMBL; AL603644; CAC49157.1; -  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 256 AA; 28459 MW; 54B064834CE7C39 CRC64;

Query Match 53.4%; Score 39; DB 16; Length 256;  
Best Local Similarity 54.5%; Pred. No. 22;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 WYXXRGIRP 12  
DB 185 WHGTRGCRPYG 195

## RESULT 8

098606 PRELIMINARY; PRT; 425 AA.  
AC 098606;  
DT 01-OCT-2001 (TREMblrel. 18, Created)  
DT 01-OCT-2001 (TREMblrel. 18, Last sequence update)  
DT 01-OCT-2001 (TREMblrel. 18, Last annotation update)  
DE ABC TRANSPORTER, SUBSTRATE BINDING PROTEIN.  
GN MLL204.  
OS Rhizobium loti (Mesorhizobium loti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Phyllobacteriaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MAFF303099;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,  
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,  
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
RA Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti.";  
RL DNA Res. 7:331-338(2000).  
DR EMBL; AP003011; BAB53357.1; -  
KW Complete proteome.  
SQ SEQUENCE 425 AA; 45011 MW; 7BF63D5600C6C177 CRC64;

Query Match 53.4%; Score 39; DB 16; Length 425;  
Best Local Similarity 60.0%; Pred. No. 38;  
Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

QY 1 AMYXXRGIRP 10  
DB 137 AMYFGRGAKP 146

## RESULT 9

000050 PRELIMINARY; PRT; 555 AA.  
AC 000050;  
DT 01-JUL-1997 (TREMblrel. 04, Created)  
DT 01-JUL-1997 (TREMblrel. 04, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE TRANSPOSASE  
OS Aspergillus niger var. awamori.  
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Eurotiomycetes;  
OC Eurotiales; Trichocomaceae; mitosporic Trichocomaceae; Aspergillus.

OX NCBI\_TaxID=105351;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-UVK143F DERIVED FROM NRRL3112; TRANSPOSON-VADER;  
 RX MEDLINE=96207472; PubMed=8625427;  
 RA Amutan M., Nyssönen E., Stubbs J., Diaz-Torres M.R., Dunn-Coleman N.;  
 RT "Identification and cloning of a mobile transposon from *Aspergillus niger* var. *awamori*."  
 RL Curr. Genet. 29:468-473(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-UVK143F DERIVED FROM NRRL3112; TRANSPOSON-VADER;  
 RX MEDLINE=97156905; PubMed=9003285;  
 RA Nyssönen E., Amutan M., Entfeld L., Stubbs J., Dunn-Coleman N.S.;  
 RT "The transposable element Tani of *Aspergillus niger* var. *awamori*, a new member of the F<sub>0</sub>I family."  
 RL Mol. Gen. Genet. 253:50-56(1996).  
 DR EMBL: U58946; AAC49623.1; -;  
 DR InterPro: IPR01878; Znf\_C2HC.  
 DR SMART: SM00343; Znf\_C2HC; 1.  
 SQ SEQUENCE 555 AA; 62978 MW; 70904D2EED09EA33 CRC64;

Query Match 53.4%; Score 39; DB 3; Length 555;  
 Best Local Similarity 50.0%; Pred. No. 52;  
 Matches 7; Conservative 1; Mismatches 6; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVGF 14  
 ||| : ||| ||  
 DB 237 AWEGOSIPTWRF 250

RESULT 10  
 O9A9U0 PRELIMINARY; PRT; 194 AA.  
 AC O9A9U0;  
 DT 01-JUN-2001 (TReMBLrel. 17, Created)  
 DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE HYPOTHETICAL PROTEIN CC0871.  
 GN CC0871.  
 OS *Caulobacter crescentus*.  
 OC Bacteria; Proteobacteria; alpha subdivision; *Caulobacter* group;  
 OC *Caulobacter*.  
 OX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19089 / CH15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nierman W.C., Feldblum T.V., Laub M.T., Paulsen I.T., Nelson K.E., Eisen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R., Potocka I., Nelson W.C., Newton A., Stephens C., Phadke N.D., Ely B., Debey R.T., Dodson R.J., Durkin A.S., Gwinn M.L., Haft D.H., Kolonay J.F., Smit J., Craven M.B., Khouri H., Shetty J., Berry K., Utterback T., Tran K., Wolf A., Vamathevan J., Ermolaeva M., White O., Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT Complete genome sequence of *Caulobacter crescentus*.  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL: AE005764; AAK22856.1; -;  
 DR TIGR: CC0871; -;  
 DR InterPro: IPR000182; Acetyltransf\_GCN5.  
 DR Pfam: PF00583; Acetyltransf\_1.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 194 AA; 21401 MW; A216C7874E8E7BE0 CRC64;

Query Match 52.1%; Score 38; DB 16; Length 194;  
 Best Local Similarity 58.3%; Pred. No. 25;  
 Matches 7; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPVG 12  
 ||| ||| |  
 DB 154 AWYERGRIRLTG 165

RESULT 11  
 O9YEH0 PRELIMINARY; PRT; 238 AA.  
 AC O9YEH0;  
 DT 01-NOV-1999 (TReMBLrel. 12, Created)  
 DT 01-NOV-1999 (TReMBLrel. 12, Last sequence update)  
 DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
 DE HYPOTHETICAL 26.1 KDA PROTEIN APE0606.  
 GN APE0606.  
 OS *Aeropyrum pernix*.  
 OC Archaea; Crenarchaeota; Desulfurococcales; Desulfurococcaceae;  
 OC *Aeropyrum*.  
 OX NCBI\_TaxID=56636;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K1;  
 RX MEDLINE=99310339; PubMed=10382966;  
 RA Kawarabayashi Y., Hino Y., Horikawa H., Yamazaki S., Haikawa Y., Jin-no K., Takahashi M., Sekine M., Baba S.-I., Ankel A., Kosugi H., Hosoyama A., Fukui S., Nagai Y., Nishijima K., Nakazawa H., Takamiya M., Masuda S., Funahashi T., Tanaka T., Kudoh Y., Yamazaki J., Kishida N., Oguchi A., Aoki K.-I., Kubota K., Nakamura Y., Nomura N., Sako Y., Kikuchi H.;  
 RT "Complete genome sequence of an aerobic hyper-thermophilic crenarchaeon, *Aeropyrum pernix* K1."  
 RL DNA Res. 6:83-101(1999).  
 DR EMBL: AP000060; BAA79576.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 238 AA; 26083 MW; CDDF1B56F04AD850 CRC64;

Query Match 52.1%; Score 38; DB 17; Length 238;  
 Best Local Similarity 63.6%; Pred. No. 31;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 1 AMYXXRGIRPV 11  
 || ||| ||  
 DB 126 AMRIRRGFRPV 136

RESULT 12  
 O9PH76 PRELIMINARY; PRT; 333 AA.  
 AC O9PH76;  
 DT 01-OCT-2000 (TReMBLrel. 15, Created)  
 DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE HYDROXYBENZATE OCTAPRENYLTRANSFERASE.  
 GN XF0068.  
 OS *Xylella fastidiosa*.  
 OC Bacteria; Proteobacteria; gamma subdivision; *Xanthomonas* group;  
 OC *Xylella*.  
 OX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M., Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S., Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Britton M.R.S., Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H., Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M., Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H., Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A., Fraga J.S., Franca S.C., Franco M.C., Fromme M., Furlan L.R., Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A., Ho P.L., Hohnselt J.D., Jungueira M.L., Kemper E.L., Kitajima J.P., Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C., Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A., Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L., Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,



RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.F.S.,  
 RA Nhani A. Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
 RA Queglio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silvestri M.L.Z., Sigleira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
 RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
 RA Zazo M.A., Zatz M., Meldalis J., Setubal J.C.,  
 RT "The genome sequence of the plant pathogen *Xylella fastidiosa*,"  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE003860; AAF8281.1; -.  
 DR InterPro: IPR000537; UblA; 1.  
 DR Pfam: PF01040; UblA; 1.  
 DR PROSITE: PS00943; UblA; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 333 AA; 37931 MW; ECF3FA716C962B95 CRC64;

Query Match 52.1%; Score 38; DB 16; Length 333;  
 Best Local Similarity 63.6%; Pred. No. 45;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

OY 2 WYXXRGIRPVG 12  
 Db 58 WKLARGDRPVG 68

RESULT 13  
 ID Q9RR92 PRELIMINARY; PRT; 398 AA.  
 AC Q9RR92;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 DE ACETATE KINASE.  
 GN DR2602.  
 OS Deinococcus radiodurans.  
 OC Bacteria; Thermus/Deinococcus group; Deinococcales; Deinococcus.  
 OX NCBI\_TaxID=1299;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=RI;  
 RX MEDLINE=20036896; PubMed=10567266;  
 RA White O., Eisen J.A., Heidelberg J.F., Hickey E.K., Peterson J.D.,  
 RA Dodson R.J., Haft D.H., Gwinn M.L., Nelson W.C., Richardson D.L.,  
 RA Moffat K.S., Qin H., Ulang L., Pamphile W., Crosby M., Shen M.,  
 RA Vanatnevan J.J., Lam P., McDonald L., Utterback T., Zaleski C.,  
 RA Makarova K.S., Aravind L., Daly M.J., Minton K.W., Fleischmann R.D.,  
 RA Kechum K.A., Nelson K.E., Salzberg S., Smith H.O., Venter J.C.,  
 RA Fraser C.M.,  
 RT "Genome sequence of the radioresistant bacterium *Deinococcus*  
 RT *radiodurans* RI,"  
 RL Science 286:1571-1577(1999).  
 DR EMBL: AE002089; AAF12139.1; -.  
 DR TIGR: DR2602; -.  
 DR InterPro: IPR000890; Acetate\_kin.  
 DR Pfam: PF00871; Acetate\_kinase; 1.  
 DR PRINTS: PR00471; ACETATEKINASE.  
 DR PROSITE: PS01076; ACETATE\_KINASE\_2; UNKNOWN\_1.  
 KW Kinase; Complete proteome.  
 SQ SEQUENCE 398 AA; 41798 MW; 79EB59FE97A15FC5 CRC64;

Query Match 52.1%; Score 38; DB 16; Length 398;  
 Best Local Similarity 50.0%; Pred. No. 55;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;  
 OY 1 AWYXXRGIRPVG 12  
 Db 111 1:1 1

Db 161 AWYXHGVRVYG 172

RESULT 14  
 ID Q95WT7 PRELIMINARY; PRT; 481 AA.  
 AC Q95WT7;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE G PROTEIN-COUPLED RECEPTOR PROTEIN (FRAGMENT).  
 GN MTH.  
 OS Drosophila yakuba (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7245;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=YAK\_30;  
 RA Duvernell D.D., Schmidt P.S., Eanes W.F.,  
 RT "Molecular population genetics of a methuselah paralog,"  
 RL submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF300415; AAK97895.1; -.  
 KW Receptor.  
 FT NON-TER  
 SQ SEQUENCE 481 AA; 55732 MW; 44773A7D5203F3FB CRC64;

Query Match 52.1%; Score 38; DB 5; Length 481;  
 Best Local Similarity 46.2%; Pred. No. 68;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 2 WYXXRGIRPVGRF 14  
 Db 297 WHNFRGTGRINRF 309

RESULT 15  
 ID Q95YN4 PRELIMINARY; PRT; 536 AA.  
 AC Q95YN4;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE G PROTEIN-COUPLED RECEPTOR.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ZIM(S)24;  
 RA Duvernell D.D., Schmidt P.S., Eanes W.F.,  
 RT "Molecular population genetics of a methuselah paralog,"  
 RL submitted (AUG-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF300396; AAK97885.1; -.  
 DR EMBL: AF300395; AAK97885.1; JOINED.  
 KW Receptor.  
 SQ SEQUENCE 536 AA; 61803 MW; E74D7E654F2E1F1 CRC64;

Query Match 52.1%; Score 38; DB 5; Length 536;  
 Best Local Similarity 46.2%; Pred. No. 77;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
 OY 2 WYXXRGIRPVGRF 14  
 Db 297 WHNFRGTGRINRF 309

Search completed: September 13, 2002, 09:29:21  
Job time: 1066 sec

---

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds  
(without alignments)  
3.891 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_8\_21

Perfect score: 69  
Sequence: 1 WYXXRGIRPVGREFX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Database : Listing first 45 summaries

1: A\_Geneseq\_032802.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT.\*  
23: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	95.7	15	21	AA49293
2	66	95.7	15	21	AA49296
3	66	95.7	20	18	AAW31394
4	66	95.7	20	20	AAW97235
5	66	95.7	20	21	AAW10365
6	66	95.7	20	21	AA49294
7	66	95.7	20	22	AA62534
8	66	95.7	20	22	AA90992
9	66	95.7	21	18	AAW31395
10	66	95.7	21	21	AAW10366
11	66	95.7	21	22	AA62535

12	66	95.7	22	18	AAW31396	Human type G prote
13	66	95.7	22	21	AAW10367	Human oxytocin sec
14	66	95.7	22	21	AA62536	Human CRH releasin
15	66	95.7	31	18	AAW31391	Human type G prote
16	66	95.7	31	20	AAW97235	Human type ligand
17	66	95.7	31	20	AAW87615	Human 19P2 ligand.
18	66	95.7	31	21	AAW10362	Human oxytocin sec
19	66	95.7	31	21	AA49291	Human CRH releasin
20	66	95.7	31	22	AA62531	19P2 ligand peptid
21	66	95.7	31	22	AA90991	Human CRH releasin
22	66	95.7	31	22	AA90995	Prolactin releasin
23	66	95.7	32	18	AAW31392	Human type G prote
24	66	95.7	32	21	AAW10363	Human oxytocin sec
25	66	95.7	32	22	AA62532	Human CRH releasin
26	66	95.7	33	18	AAW31393	Human type G prote
27	66	95.7	33	21	AAW10364	Human oxytocin sec
28	66	95.7	33	22	AA62533	Human CRH releasin
29	66	95.7	87	18	AAW31380	Human type G prote
30	66	95.7	87	20	AAW97226	Human type ligand
31	66	95.7	87	21	AAW10361	Human oxytocin sec
32	66	95.7	87	22	AA62530	Human CRH releasin
33	66	94.2	15	18	AAW31400	Synthetic ligand 1
34	65	94.2	15	20	AAW97230	C-terminal ligand
35	65	94.2	20	18	AAW31387	Rat type G protein
36	65	94.2	20	18	AAW31374	Bovine G protein-c
37	65	94.2	20	20	AAW97232	Bovine pituitary-d
38	65	94.2	20	20	AAW97234	Rat type ligand po
39	65	94.2	20	20	AAW95191	Bovine pituitary-d
40	65	94.2	20	20	AAW95175	Murine pituitary-d
41	65	94.2	20	21	AAW10350	Bovine oxytocin se
42	65	94.2	20	21	AAW10358	Rat oxytocin secre
43	65	94.2	20	21	AA49301	19P2 ligand peptid
44	65	94.2	20	21	AA49302	19P2 ligand peptid
45	65	94.2	20	22	AA62519	Bovine CRH releas

## ALIGNMENTS

RESULT	1	ALIGNMENTS
AA49293	1	
ID	AA49293 standard; peptide: 15 AA.	
XX		
AC	AA49293:	
XX		
DT	22-FEB-2000 (first entry)	
XX		
DE	19P2 ligand peptide fragment.	
XX		
KW	Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;	
KM	pituitary; regulatory mechanism; central nervous system; pancreatic.	
XX		
OS	Homo sapiens.	
XX		
FH	Key	Location/Qualifiers
FT	Modified-site 15	
FT	/note="C-terminal amide"	
XX		
PN	WO960112-A1.	
XX		
PD	25-NOV-1999.	
XX		
PF	20-MAY-1999; 99WO-JP02650.	
XX		
PR	21-MAY-1998; 98JP-0140293.	
XX		
PA	(TAKE ) TAKEDA CHEM IND LTD.	
XX		
PI	Matsumoto H, Kitada C, Hinuma S;	
XX		
DR	WPI; 2000-039381/03.	
XX		
FT	New monoclonal antibodies, useful in diagnosis, as drugs and in	

PT	studying diseases related to ligand abnormality	-
PS	Disclosure: Page 26; 73pp; Japanese.	
CC	The invention provides a monoclonal antibody which has a specific	
CC	reaction with the part peptide of the C-terminal of 19P2 ligand or its	
CC	derivative. The antibodies can be used in diagnosis or to treat or	
CC	prevent diseases associated with abnormality in the pituitary function	
CC	regulatory mechanism (e.g. promotion of prolactin secretion), central	
CC	nervous regulatory mechanism, and pancreatic function regulatory	
CC	mechanism. The antibody-based immunoassay can also be applied in	
CC	clarifying the physiological functions of the ligand and its derivative.	
CC	Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.	
SO	Sequence 15 AA:	
QY	1 WYXRGRIPVGRF 13	
DB	3 WYASRGIRPVGRF 15	
RESULT	2	
AA	AAY49296	95.7%; Score 66; DB 21; Length 15;
ID	AAY49296 standard; peptide; 15 AA.	Best Local Similarity 84.6%; Pred. No. 6.8e-06;
XX	AAY49296;	Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
XX	22-FEB-2000 (first entry)	
DT	19P2 ligand peptide fragment.	
DE	Monoclonal antibody: 19P2 ligand; diagnosis; prolactin secretion;	
XX	pituitary; regulatory mechanism; central nervous system; pancreatic.	
KW	Homo sapiens.	
XX	Location/Qualifiers	
OS	Modified-site 15	
FT	/note- "C-terminal amide"	
FT	WO9960112-A1.	
PN	25-NOV-1999.	
XX	20-MAY-1999; 99WO-JP02650.	
XX	21-MAY-1998; 98JP-0140293.	
XX	(TAKE ) TAKEDA CHEM IND LTD.	
PA	Matsumoto H, Kltada C, Hinuma S;	
XX	WPI: 2000-039381/03.	
XX	New monoclonal antibodies, useful in diagnosis, as drugs and in	
PT	studying diseases related to ligand abnormality	-
XX	Disclosure: Page 27; 73pp; Japanese.	
CC	The invention provides a monoclonal antibody which has a specific	
CC	reaction with the part peptide of the C-terminal of 19P2 ligand or its	
CC	derivative. The antibodies can be used in diagnosis or to treat or	
CC	prevent diseases associated with abnormality in the pituitary function	
CC	regulatory mechanism (e.g. promotion of prolactin secretion), central	
CC	nervous regulatory mechanism, and pancreatic function regulatory	
CC	mechanism. The antibody-based immunoassay can also be applied in	
CC	clarifying the physiological functions of the ligand and its derivative.	
CC	Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.	

XX	Sequence	15 AA:	95.7% Best Local Similarity	Score 66;	DB 21;	Length 15;
XX	Query Match		84.6%;	Pred. No. 6,8e-06;		
XX	Matches 11;	Conservative	0;	Mismatches 2;	Indels 0;	Gaps 0
OY	1 WYXXRGIRPVGRF 13					
Db	3 WYASRGIRPVGRF 15					
XX	RESULT 3					
XX	AAW31394					
XX	AAW31394 standard; Peptide; 20 AA.					
XX	AAW31394;					
XX	06-APR-1998 (first entry)					
XX	Human type G protein-coupled receptor ligand fragment 4.					
KW	G protein-coupled receptor; ligand binding; pharmaceutical;					
KM	modulator; pituitary; central nervous system; pancreas; prophylactic;					
KW	therapeutic agent.					
XX	Homo sapiens.					
OS	WO9724436-A2.					
PN	10-JUL-1997.					
PD	26-DEC-1996;	96WO-JP03821.				
PF	18-SEP-1996;	96JP-0246573.				
XX	28-DEC-1985;	95JP-0343371.				
PR	15-MAR-1986;	96JP-0059419.				
PR	12-AUG-1996;	96JP-0211805.				
XX	(TAKE ) TAKEDA CHEM IND LTD.					
PA	Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;					
PI	Kawamata Y, Kitada C;					
PI	WPI; 1997-363672/33.					
DR	N-PSDB; AAV02431.					
XX	Ligand peptide for G protein-coupled receptor - acts by modulating					
PT	function in the central nervous system, pancreas and pituitary gland					
PS	Claim 2; Page 185; 258bp; English.					
XX	This sequence represents a peptide fragment from a novel human type					
CC	ligand polypeptide corresponding to amino acid residues 34 to 53 of the					
CC	sequence represented in AAW31390 and is used in an assay to monitor					
CC	ligand binding to the G protein-coupled receptor protein. Pharmaceutical					
CC	compositions containing this ligand may be used as a pituitary function					
CC	modulator, a central nervous system modulator or a pancreatic function					
CC	prophylactic or therapeutic agent for dementia, depression, hyperkinetic					
CC	syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,					
CC	trauma, growth hormone secretory disease, hyper- and polyphagia,					
CC	hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,					
CC	hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,					
CC	Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,					
CC	transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,					
CC	acute myocardial infarction, infertility, spinocerebellar degeneration,					
CC	bone fracture, trauma, atopic dermatitis, osteoporosis and/or					
CC	oligolactia. Assays can also be developed to screen compounds which are					
CC	capable of altering the binding activity of the ligand affecting					
CC	activation of the G protein-coupled receptor protein.					

SQ Sequence 20 AA;

Query Match 95.7%; Score 66; DB 18; Length 20;  
Best Local Similarity 84.6%; Pred. No. 9.2e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 8 WYASRGIRPVGRF 20

## RESULT 4

AAM97236  
ID AAM97236 standard; peptide: 20 AA.

AC AAM97236;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

XX Rat type ligand; modulation; prolactin secretion;

KM G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;

KM Menopausal syndrome; euthyroid; hypometabolism; lactation;

KM Pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;

KM Prolactinoma; infertility; impotence; amenorrhea; galactorrhea;

KM Acromegaly; Chlari-Frommel syndrome; Argonz-del Castillo syndrome;

KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KM Contrceptive; placental function; choriocarcinoma; hydatid mole;

KM Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;

KM Abnormal lipidmetabolism; oxytocia.

XX Homo sapiens.

XX WC9856962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin

XX secretion or placental function, e.g. for treating menopausal

XX syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3; Page 166; 241pp; English.

XX The present sequence represents a human type ligand fragment. It

XX is used in the course of the invention. The specification describes

XX an agent for modulating prolactin secretion which comprises a

XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)

XX protein. The agents for promoting prolactin secretion can be used for

XX treating or preventing hypovarianism, gonocyst cacogenesis, menopausal

XX syndrome, euthyroid or hypometabolism. They can be used for promoting

XX lactation in a domestic mammal and as an aphrodisiac. The agents for

XX inhibiting prolactin secretion can be used for treating or preventing

XX pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,

XX prolactinoma, infertility, impotence, amenorrhea, galactorrhea,

XX acromegaly, Chlari-Frommel syndrome, Argonz-del Castillo syndrome,

XX Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.

SQ Sequence 20 AA;

Query Match 95.7%; Score 66; DB 20; Length 20;  
Best Local Similarity 84.6%; Pred. No. 9.2e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 8 WYASRGIRPVGRF 20

## RESULT 5

AAB10365  
ID AAB10365 standard; peptide: 20 AA.

AC AAB10365;

DT 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.

XX Human: oxytocin secretion promoter; G protein-coupled receptor protein;

KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;

KM caesarean section; artificial fertilization; galactostasis; goat; pig;

KM veterinary medicine; milk production.

XX Homo sapiens.

XX WO200038704-A1.

XX 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP07199.

XX 25-DEC-1998; 98JP-0369585.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kikada C, Hinuma S;

XX WPI; 2000-452298/39.

XX Physiologically-active polypeptide recognized as ligand by G

XX protein-coupled receptor protein, for promoting secretion of oxytocin,

XX as drugs for diseases relating to oxytocin secretion and in veterinary

XX medicine

XX Disclosure; Page 63; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent

XX which contains a ligand peptide or its salt for the G protein-coupled

XX receptor protein. It is useful in the form of drugs for ameliorating,

XX preventing and treating diseases relating to oxytocin secretion e.g.

XX weak pains and atonic bleeding, before and after expulsion of placenta,

XX uterine recovery failure, caesarean section, stoppage of artificial

XX fertilization or galactostasis and is also applicable in veterinary

XX medicine for promoting milk production in cow, goat and pig. This

XX sequence represents a human peptide which acts as an oxytocin secretion

XX promoter.

XX Sequence 20 AA;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 8 WYASRGIRPVGRF 20

RESULT 6  
AAV49294  
ID AAV49294 standard; peptide; 20 AA.  
XX  
AC AAV49294;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE 19P2 ligand peptide fragment.  
XX  
DE 19P2 ligand peptide fragment.  
XX  
KM Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
OS Homo sapiens.  
XX  
FH Key Location/Qualifiers  
FT Modified-site 20  
FT /note="C-terminal amide"  
XX  
PN WO960112-A1.  
XX  
PD 25-NOV-1999.  
XX  
PF 20-MAY-1999; 99WO-JP02650.  
XX  
PR 21-MAY-1998; 98JP-0140293.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
DR WPI: 2000-039381/03.  
XX  
PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
PS Disclosure: Page 26; 73pp; Japanese.  
XX  
XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAV49290-302 represent peptide fragments of the 19P2 ligand.  
XX  
SQ Sequence 20 AA:  
  
Query Match 95.7%; Score 66; DB 21; Length 20;  
Best Local Similarity 84.6%; Pred. No. 9.2e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 WYXXRGIRPYGRF 13  
DB 8 WYASRGIRPYGRF 20  
  
RESULT 7  
AAG62534  
ID AAG62534 standard; peptide; 20 AA.  
XX  
AC AAG62534;  
XX  
DT 24-NOV-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 35.  
XX  
KW Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.

XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000WO-JP08119.  
XX  
PR 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Kitada C, Matsumoto H, Hinuma S;  
DR WPI: 2001-355552/37.  
XX  
PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Claim 4; Page 75; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 20 AA:  
  
Query Match 95.7%; Score 66; DB 22; Length 20;  
Best Local Similarity 84.6%; Pred. No. 9.2e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 WYXXRGIRPYGRF 13  
DB 8 WYASRGIRPYGRF 20  
  
RESULT 8  
AAB90992  
ID AAB90992 standard; peptide; 20 AA.  
XX  
AC AAB90992;  
XX  
DT 22-JUN-2001 (first entry)  
XX  
DE Prolactin releasing peptide SEQ ID NO:166.  
XX  
KW Protection; endogenous therapeutic peptide; peptidase; conjugation;  
KW blood component; modification; succinimidyl; maleimido group; amino;  
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.  
XX  
OS Homo sapiens.  
OS Synthetic.  
XX  
PN WO200069900-A2.  
XX  
PD 23-NOV-2000.  
XX  
PF 17-MAY-2000; 2000WO-US13576.  
XX  
PR 17-MAY-1999; 99US-0134406.  
PR 10-SEP-1999; 99US-0153406.  
PR 15-OCT-1999; 99US-0159783.  
XX

PA (CONU-) CONNUCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
PS  
XX Disclosure; Page 244; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidase stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specificity as bonding to large molecules decreases  
CC intracellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 20 AA:  
  
Query Match 95.7%; Score 66; DB 22; Length 20;  
Best Local Similarity 84.6%; Pred. No. 9.2e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 WYXXRGIRPVGRF 13  
II |  
Db 8 WYASRGIRPVGRF 20  
  
RESULT 9  
AABW31395  
ID AABW31395 standard; peptide; 21 AA.  
XX  
AC AABW31395;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 5.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
PI Kawamata Y, Kitada C;

XX  
DR WPI; 1997-363672/33.  
DR N-PSDB; AAV02432.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
PS  
XX Claim 2; Page 186; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AABW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphegia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 21 AA:  
  
Query Match 95.7%; Score 66; DB 18; Length 21;  
Best Local Similarity 84.6%; Pred. No. 9.7e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 WYXXRGIRPVGRF 13  
II |  
Db 8 WYASRGIRPVGRF 20  
  
RESULT 10  
AAB10366  
ID AAB10366 standard; peptide; 21 AA.  
XX  
AC AAB10366;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
XX  
DE Human oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO2000038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07189.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-452298/39.  
PT Physiologically-active polypeptide recognized as ligand by G

PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
 as drugs for diseases relating to oxytocin secretion and in veterinary  
 medicine -  
 XX  
 XX  
 PS Disclosure; Page 63; 72pp; Japanese.

CC This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a human peptide which acts as an oxytocin secretion  
 CC promoter.

SO Sequence 21 AA;

Query Match 95.7%; Score 66; DB 21; Length 21;  
 Best Local Similarity 84.6%; Pred. No. 9.7e-06;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
 || |||||  
 Db 8 WYASRGIRPVGRF 20

RESULT 11

AAG62535  
 ID AAG62535 standard; peptide; 21 AA.

XX AAG62535;

DT 24-AUG-2001 (first entry)

DE Human CMH releasing protein related peptide SEQ ID NO: 36.

XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
 KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
 KM Addison's disease; adrenal gland hyperfunction; obesity.

XX Homo sapiens.

XX WO200135984-A1.

PD 25-MAY-2001.

PF 17-NOV-2000; 2000WO-JP08119.

PR 18-NOV-1999; 99JP-0327900.

PR 26-SEP-2000; 2000JP-0297073.

PA (TAKE ) TAKEDA CHEM IND LTD.

XX Kitada C, Matsumoto H, Hinuma S;

DR WPI: 2001-355552/37.

PT Use of G protein receptor ligand or peptide for controlling  
 PT corticotrophin releasing hormone secretion -

XX Disclosure; Page 75; 90pp; Japanese.

CC The present sequence describes a method of controlling the secretion of  
 CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
 CC receptor ligand. This can be used to control the secretion of CRH and is  
 CC useful as an analgesic or for treating, preventing or ameliorating  
 CC diseases associated with CRH secretion such as hyperaldosteronism,  
 CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
 CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
 CC loss, and hypotension), adrenal gland hypofunction and obesity. The

CC Present sequence is a peptide used in the exemplification of the  
 CC invention.

SO Sequence 21 AA;

Query Match 95.7%; Score 66; DB 22; Length 21;  
 Best Local Similarity 84.6%; Pred. No. 9.7e-06;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
 || |||||  
 Db 8 WYASRGIRPVGRF 20

RESULT 12

AAW31396  
 ID AAW31396 standard; peptide; 22 AA.

XX AAW31396;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 6.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
 KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KM therapeutic agent.

XX Homo sapiens.

XX WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX WPI: 1997-363672/33.

DR N-PSDB; AAW02433.

PS Claim 2; Page 186; 258pp; English.

CC This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
 CC sequence represented in AAW31390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidemia, hypercholesterolemia, hyperglyceridemia,  
 CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neuritis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting



CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 22 AA;

Query Match 95.7%; Score 66; DB 18; Length 22;  
Best Local Similarity 84.6%; Pred. No. 1e-05;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 8 WYASRGIRPVGRF 20

RESULT 13  
AAB10367  
ID AAB10367 standard; peptide: 22 AA.  
XX  
AC AAB10367;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.  
XX

KM Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX

PD 06-JUL-2000.  
XX  
PE 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI: 2000-45298/39.  
XX

PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
PS Disclosure: Page 64; 72pp; Japanese.  
XX

CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
SQ Sequence 22 AA;

Query Match 95.7%; Score 66; DB 21; Length 22;  
Best Local Similarity 84.6%; Pred. No. 1e-05;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 8 WYASRGIRPVGRF 20

RESULT 14  
AAG62536  
ID AAG62536 standard; peptide: 22 AA.  
XX  
AC AAG62536;  
XX

DT 24-AUG-2001 (first entry)  
XX

DE Human CRH releasing protein related peptide SEQ ID NO: 37.  
XX

KM Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
OS Homo sapiens.  
XX  
PN WO200135984-A1.  
XX

PD 25-MAY-2001.  
XX

PE 17-NOV-2000; 2000WO-JP08119.  
XX

PR 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Kitada C, Matsumoto H, Hinuma S;  
XX  
DR WPI: 2001-35552/37.  
XX

PT Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
PS Disclosure: Page 75; 90pp; Japanese.  
XX

CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
SQ Sequence 22 AA;

Query Match 95.7%; Score 66; DB 22; Length 22;  
Best Local Similarity 84.6%; Pred. No. 1e-05;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 8 WYASRGIRPVGRF 20

RESULT 15  
AAW31391  
ID AAW31391 standard; Peptide: 31 AA.  
XX  
AC AAW31391;  
XX

DT 06-APR-1998 (first entry)  
XX

DE Human type G protein-coupled receptor ligand fragment 1.  
XX

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;

KW therapeutic agent.

**Homo sapiens.**

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 15-MAR-1996; 96JP-0059419.

[illegible][illegible]

PI Kawamata Y, Kltada C;  
vv

DR WPI; 1997-363672/33.  
DR N-PCDD; 199702430

**XX**

function in the central nervous system, pancreas and pituitary gland

PS Claim 2; Page 184; 258pp; English.  
XY

This sequence represents a peptide fragment from a novel human type I band polypeptide corresponding to the amino acid residues 33 to 73 of the

CC sequence represented in AAW1330 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmacological  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and hypophagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
CC hyperproliferidaemia, diabetes, cancer, pancreaticidaemia, disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.

**SQ**      **Sequence**      **31 AA;**

Query Match	Score	DB	Length
95.78;	66;	18;	31;

```
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

QY 1 WYXXRGIRPVGRF 13

Db 19 wystrglrpvgrf 31

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 seconds  
(without alignments)  
2.503 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_8\_21

Perfect score: 69

Sequence: 1 WYXXRCIRPVGRFX 14

Scoring table: BLOSUM62

Searched: Gapop 10.0 , Gapext 0.5

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCtus\_COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	95.7	20	3	US-09-105-678A-46
2	66	95.7	20	4	US-08-776-971-64
3	66	95.7	20	4	US-09-421-208-46
4	66	95.7	21	3	US-09-105-678A-47
5	66	95.7	21	4	US-08-776-971-65
6	66	95.7	21	4	US-09-421-208-47
7	66	95.7	22	3	US-09-105-678A-48
8	66	95.7	22	4	US-08-776-971-66
9	66	95.7	22	4	US-09-421-208-48
10	66	95.7	31	3	US-09-105-678A-9
11	66	95.7	31	4	US-08-776-971-61
12	66	95.7	31	4	US-09-421-208-9
13	66	95.7	31	4	US-09-105-678A-43
14	66	95.7	31	4	US-09-421-208-43
15	66	95.7	32	3	US-09-105-678A-44
16	66	95.7	32	4	US-08-776-971-62
17	66	95.7	32	4	US-09-421-208-44
18	66	95.7	33	3	US-09-105-678A-45
19	66	95.7	33	4	US-08-776-971-63
20	66	95.7	33	4	US-09-421-208-45
21	66	95.7	37	4	US-08-776-971-59
22	66	95.7	87	4	US-08-776-971-135
23	66	95.7	87	4	US-08-776-971-138
24	65	94.2	15	4	US-08-776-971-93
25	65	94.2	20	3	US-09-105-678A-34
26	65	94.2	20	3	US-09-105-678A-40
27	65	94.2	20	4	US-08-776-971-8

28	65	94.2	20	4	US-08-776-971-50	Sequence 50, Appl
29	65	94.2	20	4	US-08-776-971-98	Sequence 98, Appl
30	65	94.2	20	4	US-09-421-208-34	Sequence 34, Appl
31	65	94.2	20	4	US-09-421-208-40	Sequence 40, Appl
32	65	94.2	21	3	US-09-105-678A-35	Sequence 35, Appl
33	65	94.2	21	3	US-09-105-678A-41	Sequence 41, Appl
34	65	94.2	21	4	US-08-776-971-9	Sequence 9, Appl1
35	65	94.2	21	4	US-08-776-971-51	Sequence 51, Appl1
36	65	94.2	21	4	US-09-421-208-35	Sequence 35, Appl
37	65	94.2	21	4	US-09-421-208-41	Sequence 41, Appl
38	65	94.2	22	3	US-09-105-678A-36	Sequence 36, Appl
39	65	94.2	22	3	US-09-105-678A-42	Sequence 42, Appl
40	65	94.2	22	4	US-08-776-971-10	Sequence 10, Appl
41	65	94.2	22	4	US-08-776-971-52	Sequence 52, Appl
42	65	94.2	22	4	US-09-421-208-36	Sequence 36, Appl
43	65	94.2	22	4	US-09-421-208-42	Sequence 42, Appl
44	65	94.2	31	3	US-09-105-678A-7	Sequence 7, Appl1
45	65	94.2	31	3	US-09-105-678A-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
; Sequence 46, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19p2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESS: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA: US/09/105, 678A  
; APPLICATION NUMBER: US/09/105, 678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-46

Query Match 95.7%; Score 66; DB 3; Length 20;  
Best Local Similarity 84.6%; Pred. No. 2.7e+06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0;

OY 1 WYXXRGIRPVGRF 13  
11 |  
Db 8 WYASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64  
; Sequence 64, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 95.7%; Score 66; DB 4; Length 20;  
Best Local Similarity 84.6%; Pred. No. 2.7e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
OY 1 WYXXRGIRPVGRF 13  
11 |  
Db 8 WYASRGIRPVGRF 20

## RESULT 3

US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 95.7%; Score 66; DB 4; Length 20;  
Best Local Similarity 84.6%; Pred. No. 2.7e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
11 |  
Db 8 WYASRGIRPVGRF 20

## RESULT 4

US-09-105-678A-47  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston

STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 95.7%; Score 66; DB 3; Length 21;  
Best Local Similarity 84.6%; Pred. No. 2.8e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
DB 8 WYASRGIRPVGRF 20

RESULT 5  
US-08-776-971-65  
Sequence 65, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996

APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 65:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 65:  
US-08-776-971-65

Query Match 95.7%; Score 66; DB 4; Length 21;  
Best Local Similarity 84.6%; Pred. No. 2.8e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
DB 8 WYASRGIRPVGRF 20

RESULT 6  
US-09-421-208-47  
Sequence 47, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 47:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 21 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 95.7%; Score 66; DB 4; Length 21;  
Best Local Similarity 84.6%; Pred. No. 2.9e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXRGIRPVGRF 13  
11 |||||||  
DB 8 WYASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
; Sequence 48, Application US/09105678A  
; Patent No. 6103882

GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 48:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 95.7%; Score 66; DB 3; Length 22;  
Best Local Similarity 84.6%; Pred. No. 2.9e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXRGIRPVGRF 13  
11 |||||||  
DB 8 WYASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
; Sequence 66, Application US/08776971B  
; Patent No. 6228984

GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 66:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 22 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 95.7%; Score 66; DB 4; Length 22;  
Best Local Similarity 84.6%; Pred. No. 2.9e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXRGIRPVGRF 13  
11 |||||||  
DB 8 WYASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

```
; Sequence 48, Application US/09421208
; Patent No. 6238561
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/421,208
; FILING DATE:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 09/105,678
; FILING DATE: 26-JUN-1998
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-6440
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 48:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 22 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-421-208-48

Query Match          95.7%; Score 66; DB 4; Length 22;
Best Local Similarity 84.6%; Pred. No. 2.9e-06;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 31 amino acids
; TYPE: amino acid
; STRANDEDNESS:
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; US-09-105-678A-9
```

```
Query Match          95.7%; Score 66; DB 3; Length 31;
Best Local Similarity 84.6%; Pred. No. 4.3e-06;
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
```

```
OY      1 WYXXRGIRPVGRF 13
      || |||||
Db      19 WYASRGIRPVGRF 31

RESULT 11
; Sequence 43, Application US/09105678A
; Patent No. 6103882
; GENERAL INFORMATION:
; APPLICANT: Suenaga, Masato
; APPLICANT: Moriya, Takeo
; APPLICANT: Tanaka, Yoko
; APPLICANT: Nishimura, Osamu
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
; NUMBER OF SEQUENCES: 52
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
; STREET: 130 Water Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/105,678A
; FILING DATE: 26-JUN-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 172118/1997
; FILING DATE: 27-JUN-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Conlin, David G.
; REGISTRATION NUMBER: 27,026
; REFERENCE/DOCKET NUMBER: 48466-342
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-523-3400
; TELEFAX: 617-523-6440
```

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-43

Query Match  
Best Local Similarity 95.7%; Score 66; DB 3; Length 31;  
Best Local Similarity 84.6%; Pred. No. 4.3e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPYGRF 13  
|| |||||  
Db 19 WYASRGIRPYGRF 31

RESULT 12  
US-08-776-971-61  
Sequence 61, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitade, Chiko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 61:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match  
Best Local Similarity 95.7%; Score 66; DB 4; Length 31;  
Best Local Similarity 84.6%; Pred. No. 4.3e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPYGRF 13  
|| |||||  
Db 19 WYASRGIRPYGRF 31

RESULT 13  
US-09-421-208-9  
Sequence 9, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
Moriya, Takeo  
Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421, 208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105, 678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match  
Best Local Similarity 95.7%; Score 66; DB 4; Length 31;  
Best Local Similarity 84.6%; Pred. No. 4.3e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPYGRF 13  
|| |||||  
Db 19 WYASRGIRPYGRF 31

RESULT 14



US-09-421-208-43  
; Sequence 43, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 43:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 31 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-421-208-43

Query Match 95.7%; Score 66; DB 4; Length 31;  
Best Local Similarity 84.6%; Pred. No. 4.3e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 19 WYASRGIRPVGRF 31

RESULT 15  
US-09-105-678A-44  
; Sequence 44, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 44:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 32 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
; US-09-105-678A-44

Query Match 95.7%; Score 66; DB 3; Length 32;  
Best Local Similarity 84.6%; Pred. No. 4.4e-06;  
Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
|| |||||  
Db 19 WYASRGIRPVGRF 31

Search completed: September 13, 2002, 09:20:58  
job time: 623 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:01 ; Search time 172.41 Seconds  
(without alignments)  
7.803 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_8\_21

Perfect score: 69

Sequence: 1 WYXXRGIRPVGRFX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR.71.\*  
2: PIR1.\*  
3: PIR2.\*  
4: PIR3.\*  
5: PIR4.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	65	94.2	83	2	JC7607
2	42	60.9	719	2	S61046
3	39	56.5	256	2	E95936
4	39	56.5	485	2	E83663
5	38	55.1	333	2	H82852
6	38	55.1	2870	2	A35458
7	37.5	54.3	154	2	AH0264
8	37	53.6	149	1	NEBP37
9	37	53.6	338	2	D95851
10	37	53.6	343	2	E95986
11	37	53.6	395	2	T34570
12	37	53.6	769	2	T51112
13	37	53.6	788	2	I64045
14	36	52.2	142	2	S60747
15	36	52.2	184	2	T35841
16	36	52.2	276	2	D70817
17	36	52.2	293	2	C61408
18	36	52.2	293	2	G97409
19	36	52.2	293	2	AH2627
20	36	52.2	338	2	T20100
21	36	52.2	488	2	B89957
22	36	52.2	503	2	A82193
23	36	52.2	545	2	A67448
24	36	52.2	674	2	G70875
25	36	52.2	790	2	T47959
26	36	52.2	966	2	S43878
27	36	52.2	975	2	T30816
28	35	50.7	68	2	D72428
29	35	50.7	105	2	D81319

30	35	50.7	115	2	G72674	hypothetical prote
31	35	50.7	178	2	D95026	ribosomal protein
32	35	50.7	178	2	D97897	50S ribosomal prot
33	35	50.7	269	2	T44866	hypothetical prote
34	35	50.7	272	2	T39533	hypothetical prote
35	35	50.7	322	2	D71686	hypothetical prote
36	35	50.7	326	2	A97751	hypothetical prote
37	35	50.7	391	1	XXG2AC	acetyl-CoA C-acety
38	35	50.7	453	2	S18597	tubulin beta chain
39	35	50.7	460	2	D84396	signal recognition
40	35	50.7	504	2	B70821	hypothetical prote
41	35	50.7	519	2	T45447	probable two-compo
42	35	50.7	566	2	A84562	probable integral
43	35	50.7	618	2	S06446	citrolysin-related
44	35	50.7	894	2	E82221	DNA gyrase, chain
45	35	50.7	1056	2	B82557	hypothetical prote

#### ALIGNMENTS

```
RESULT 1
JC7607
prolactin-releasing peptide - rat
C:Species: Rattus norvegicus (Norway rat)
C:Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 30-Jun-2001
C:Accession: JC7607
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya,
Biochem. Biophys. Res. Commun. 281, 53-56, 2001
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:
A:Reference number: JC7607; MUID:21092785; PMID:11178959
A:Contents: Spleen
A:Accession: JC7607
A:Molecule type: DNA
A:Residues: 1-83 <YAM>
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior
release, and stimulation of ACTH secretion from the pituitary.
C:Genetics:
A:Gene: PrRP
A:Introns: 33/1

Query Match          94.2%  Score 65;  DB 2;  Length 83;
Best Local Similarity 84.6%  Pred. 4.3e-05;
Matches 11;  Conservative 0;  Mismatches 2;  Indels 0;  Gaps 0;

OY      1 WYXXRGIRPVGRF 13
      11 | | | | | | | |
DB      40 WYTGRCIRPVGRF 52

RESULT 2
S61046
A:Protein - yeast (Saccharomyces cerevisiae)
A:Alternate names: protein D1478; protein YDL167c
C:Species: Saccharomyces cerevisiae
C:Date: 15-Feb-1996 #sequence_revision 01-Mar-1996 #text_change 29-Oct-1999
C:Accession: S61046; S31139; S67719
R:Polh, T.M.
submitted to the EMBL Data Library, November 1995
A:Reference number: S61010
A:Accession: S61046
A:Molecule type: DNA
A:Residues: 1-719 <POH>
A:Cross-references: EMBL:Z67750; NID:q1061256; PIDN:CAA91579.1; PID:q1061272
R:Wehner, E.P.; Rao, E.; Brendel, M.
Mol. Gen. Genet. 237, 351-358, 1993
A:Title: Molecular structure and genetic regulation of SFA, a gene responsible for re
A:Reference number: S31138; MUID:93247548
A:Accession: S31139
A>Status: nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
```

A:Residues: 1-492, 'N', 494-719 <MEH>  
A:Cross-references: EMBL:K68020; NID:9577609; PIDN:CAA48159.1; PID:9288590  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, August 1992  
R:Polh, T.M.  
Submitted to the Protein Sequence Database, July 1996  
A:Reference number: S67708  
A:Accession: S67719  
A:Molecule type: DNA  
A:Residues: 1-719 <POM>  
A:Cross-references: EMBL:Z74215; NID:9431265; PIDN:CAA9741.1; PID:e253076; PID:g143126  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: SGD:MRP1, ARP1  
A:Cross-references: MIPS:YDL167c; SGD:S0002326  
A:Map position: 4L

Query Match 60.9%; Score 42; DB 2; Length 719;  
Best Local Similarity 54.5%; Pred. No. 7.8;  
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 1 WYXXRGIRPYG 11  
| : |||||  
Db 245 WFTQYGVPRYG 255

RESULT 3  
E95936  
hypothetical protein [imported] - Sinorhizobium meliloti (strain 1021) megaplasmid pSymE  
C:Species: Sinorhizobium meliloti  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: E95936  
R:Finan, T.M.; Welder, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Hernan  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing endo  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: E95936  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-256 <NUR>  
A:Cross-references: GB:AL591985; PIDN:CAC49157.1; PID:g15140642; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid pSymB  
R:Calbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,  
P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kles, E.; Komp, C.; LeLaure,  
hepault, P.; Vandenbol, M.; Vorholter, F.J.; Weldner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMD21253  
A:Genome: plasmid

Query Match 56.5%; Score 39; DB 2; Length 256;  
Best Local Similarity 54.5%; Pred. No. 10;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 1 WYXXRGIRPYG 11  
| : |||||  
Db 185 WHGTGCRPYG 195

RESULT 4  
E83663  
glutamyI-tRNA synthetase gltX [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: E83663  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
Nucleic Acids Res. 28, 4317-4331, 2000

A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: E83663  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-485 <STO>  
A:Cross-references: GB:AP001507; GB:BA000004; NID:g10172612; PIDN:BA03828.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: gltX  
C:Superfamily: glutamate--tRNA ligase; glutamine--tRNA ligase homology

Query Match 56.5%; Score 39; DB 2; Length 485;  
Best Local Similarity 58.3%; Pred. No. 19;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 2 YXXRGIRPYGRF 13  
| : |||||  
Db 146 YEAKGIRPYGRF 157

RESULT 5  
H82852  
hydroxybenzoate octaprenyltransferase XF0068 [imported] - Xylella fastidiosa (strain  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 17-Nov-2000  
C:Accession: H82852  
R:Anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: H82852  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-333 <SIM>  
A:Cross-references: GB:AE003860; GB:AE003849; NID:g9104830; PIDN:AAF82881.1; GSPDB:GN  
A:Experimental source: strain 945c  
R:Simpson, A.J.G.; Relnach, F.C.; Arrude, P.; Abreu, F.A.; Acencio, M.; Alverenga, R.  
Briones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer  
as-Neco, E.; Docena, C.; El-Dorri, H.; Pacincaul, A.B.; Ferreira, A.V.S.  
Submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, C.  
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Sanceli, R.V.; Sawa  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovskij-Almeida, S.; Vettore, A.L.  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF0068  
C:Superfamily: 4-hydroxybenzoate octaprenyltransferase

Query Match 55.1%; Score 38; DB 2; Length 333;  
Best Local Similarity 63.6%; Pred. No. 20;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WYXXRGIRPYG 11  
| : |||||  
Db 58 WKLRGDRPYG 68

RESULT 6  
A35548  
319k protein ndvB - Rhizobium meliloti  
C:Species: Rhizobium meliloti  
C>Date: 21-Sep-1990 #sequence\_revision 21-Sep-1990 #text\_change 08-Oct-1999  
C:Accession: A35548

R:Jalpi, L.; Dylan, T.; Ditta, G.S.; Helinski, D.R.; Stanfield, S.W.  
J. Biol. Chem. 265, 2843-2851, 1990  
A:Title: The ndvB locus of *Rhizobium meliloti* encodes a 319-kDa protein involved in the  
A:Reference number: A35548; MUID:90153914  
A:Accession: A35548  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-2870 <TEL>  
A:Cross-references: GB:J05219; NID:g152270; PIDN:AAA26305.1; PID:g152271  
C:Keywords: transmembrane protein

Query Match 55.18; Score 38; DB 2; Length 2870;  
Best Local Similarity 50.0%; Pred. No. 1.7e+02;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 WYXXRGIRPV 10  
DB 1126 WYRTGLEPM 1135

RESULT 7  
AH0264  
conserved hypothetical protein YP02172 [imported] - *Yersinia pestis* (strain CO92)  
C:Species: *Yersinia pestis*  
C:Date: 02-Nov-2001 #sequence\_revision 02-Nov-2001 #text\_change 09-Nov-2001  
C:Accession: AH0264  
R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Tlthbl, R.W.; Holden, M.T.G.; Prentice, M.B.  
deno-Tarraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Davis, P.; Dougan, G.;  
11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,  
Nature 413, 533-527, 2001  
A:Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.  
A:Reference number: AB0001; MUID:21470413; PMID:11586360  
A:Accession: AH0264  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-154 <KUR>  
A:Cross-references: GB:AL590842; PIDN:CAC90980.1; PID:g15980175; GSPDB:GN00175  
C:Genetics:  
A:Gene: YP02172  
C:Superfamily: hypothetical protein HI0277

Query Match 54.38; Score 37.5; DB 2; Length 154;  
Best Local Similarity 61.5%; Pred. No. 12;  
Matches 8; Conservative 1; Mismatches 3; Indels 1; Gaps 1;

OY 1 WYXXRGIRPVGR 12  
DB 124 WYRIDGIRPSLGR 136

RESULT 8  
NEBP37  
endodeoxyribonuclease I (EC 3.1.21.-) - phage T7  
C:Species: phage T7  
C:Date: 01-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 11-Jun-1999  
C:Accession: B94615; C92866; S42301; A00785  
R:Dunn, J.J.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A:Reference number: A94615  
A:Accession: B94615  
A:Molecule type: DNA  
A:Residues: 1-149 <DUI>  
R:Dunn, J.J.; Studier, F.W.  
J. Mol. Biol. 148, 303-330, 1981  
A:Title: Nucleotide sequence from the genetic left end of bacteriophage T7 DNA to the be  
A:Reference number: A92866; MUID:82078034  
A:Accession: C92866  
A:Molecule type: DNA  
A:Residues: 1-149 <DUI>  
A:Cross-references: GB:Y01127; NID:g15496; PIDN:CAA24345.1; PID:g15517  
R:Dunn, J.J.; Studier, F.W.

J. Mol. Biol. 166, 477-535, 1983  
A:Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7  
A:Reference number: S42283; MUID:83241725  
A:Accession: S42301  
A:Molecule type: DNA  
A:Residues: 1-149 <DUI>  
A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24402.1; PID:g15581  
A:Note: The authors did not translate the codon for residue 1  
C:Comment: Endodeoxyribonuclease I, which is expressed in the late stage, is necessary  
s a linear monomer. In the late stage, the T7 DNA replicates via linear concatamers s  
C:Genetics:  
A:Gene: 3  
A:Map position: 25.64-26.76  
C:Superfamily: phage T7 endodeoxyribonuclease I  
C:Keywords: hydrolase

Query Match 53.68; Score 37; DB 1; Length 149;  
Best Local Similarity 58.3%; Pred. No. 14;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 YXXRGIRPVGRF 13  
DB 4 YGANGIRRVGAF 15

RESULT 9  
D95851  
Probable oxidoreductase protein Smb20076 [imported] - *Sinorhizobium meliloti* (strain  
C:Species: *Sinorhizobium meliloti*  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: D95851  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-fixing e  
A:Reference number: A95842; MUID:2136508; PMID:11481431  
A:Accession: D95851  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-338 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAC48476.1; PID:g15139948; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid pSymB  
R:Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub  
pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau  
nebault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: Smb20076  
A:Genome: Plasmid

Query Match 53.68; Score 37; DB 2; Length 338;  
Best Local Similarity 46.2%; Pred. No. 32;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
DB 87 WFSNVGAVGAVGRF 99

RESULT 10  
E95986  
conserved hypothetical membrane protein Smb20863 [imported] - *Sinorhizobium meliloti*  
C:Species: *Sinorhizobium meliloti*  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: E95986  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001

A:Title: The complete sequence of the 1,683-kb pSymb megaplasmid from the N2-fixing endo  
 A:Reference number: A95842; MUID:21396508; PMID:11481431  
 A:Accession: E95986  
 A:Status: preliminary  
 A:Molecule type: DNA  
 A:Residues: 1-343 <KUR>  
 A:Cross-references: GB:AL591985; PIDN:CA049557.1; PID:g15141044; GSPDB:GN00167  
 A:Experimental source: strain 1021, megaplasmid pSymb  
 R:Galbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Bartley-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.  
 Science 293, 668-672, 2001  
 A:Authors: Kahn, D.; Kahn, M.L.; Kallman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure, H.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.  
 A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
 A:Reference number: A96039; MUID:21368234; PMID:11474104  
 A:Contents: annotation  
 C:Genetics:  
 A:Gene: SMD20863  
 A:Genome: plasmid

Query Match 53.6%; Score 37; DB 2; Length 343;  
 Best Local Similarity 50.0%; Pred. No. 32;  
 Matches 6; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

OY 1 WYXXRCIRPVGR 12  
 | : | | | |  
 DB 24 WHIOGCTRPTRGR 35

RESULT 11  
 T34570  
 probable oxidase - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T34570  
 R:Murphy, L.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 submitted to the EMBL Data Library, January 1998  
 A:Reference number: Z21548  
 A:Accession: T34570  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-395 <KUR>  
 A:Cross-references: EMBL:AL021529; PIDN:CA16434.1; GSPDB:GN00070; SCOEDB:SC10A5.02  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC10A5.02

Query Match 53.6%; Score 37; DB 2; Length 395;  
 Best Local Similarity 46.2%; Pred. No. 37;  
 Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 WYXXRCIRPVGR 13  
 | : | | | |  
 DB 269 WYEQVKPPLRP 281

RESULT 12  
 T51112  
 beta-glucosidase (EC 3.2.1.21) oler [validated] - Streptomyces antibioticus (ATCC 11891)  
 C:Species: Streptomyces antibioticus  
 A:Variety: ATCC 11891  
 C:Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 21-Jul-2000  
 C:Accession: T51112  
 R:Quirós, L.M.; Aguilrazabalaga, I.; Olano, C.; Mendez, C.; Salas, J.A.  
 submitted to the EMBL Data Library, September 1999  
 A:Reference number: Z25300  
 A:Accession: T51112  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1-769 <OUT>

A:Cross-references: EMBL:AF055579; PIDN:ANCI2650.1  
 A:Experimental source: ATCC 11891  
 C:Genetics:  
 A:Gene: oler  
 C:Function:  
 A:Description: extracellular reactivation of oleandomycin; converts glycosylated olea  
 C:Keywords: glycosidase; hydrolase

Query Match 53.6%; Score 37; DB 2; Length 769;  
 Best Local Similarity 38.5%; Pred. No. 72;  
 Matches 5; Conservative 4; Mismatches 4; Indels 0; Gaps 0;

OY 1 WYXXRCIRPVGR 13  
 | : | | | |  
 DB 637 WYEQGVQPLPFSF 649

RESULT 13  
 I64045  
 recombinant protein rec2 - Haemophilus influenzae (strain Rd KW20)  
 C:Species: Haemophilus influenzae  
 C:Date: 18-Aug-1995 #sequence\_revision 18-Aug-1995 #text\_change 08-Oct-1999  
 C:Accession: I64045  
 R:Flieschmann, R.D.; Adams, M.D.; White, O.; Clayton, R.A.; Kirkness, E.F.; Kerlavage  
 ; Gocayne, J.D.; Scott, J.; Shirley, R.; Liu, L.I.; Glodek, A.; Kelley, J.M.; Weidman  
 ; D.M.; Brandon, R.C.; Fine, L.D.; Fritchman, J.L.; Fuhrmann, J.L.; Geoghegan, N.S.M.  
 Science 269, 496-512, 1995  
 A:Authors: Gnehm, C.L.; McDonald, L.A.; Small, K.V.; Fraser, C.M.; Smith, H.O.; Venter  
 A:Title: Whole-genome random sequencing and assembly of Haemophilus influenzae Rd.  
 A:Reference number: A64000; MUID:95350630  
 A:Accession: I64045  
 A:Status: nucleic acid sequence not shown; translation not shown  
 A:Molecule type: DNA  
 A:Residues: 1-788 <TIGR>  
 A:Cross-references: GB:U032691; GB:LA2023; NID:g1573004; PIDN:ANCI21739.1; PID:g1573009  
 C:Genetics:  
 A:Gene: rec2

Query Match 53.6%; Score 37; DB 2; Length 788;  
 Best Local Similarity 54.5%; Pred. No. 74;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 WYXXRCIRPVGR 11  
 | : | | | |  
 DB 156 WYFSKGTAVG 166

RESULT 14  
 S60747  
 bacitracin synthetase - Bacillus licheniformis (fragment)  
 C:Species: Bacillus licheniformis  
 C:Date: 27-Apr-1996 #sequence\_revision 13-Mar-1997 #text\_change 03-Nov-2000  
 C:Accession: S60747  
 R:Turgey, K.; Marahel, M.A.  
 Pept. Res. 7, 238-241, 1994  
 A:Title: A general approach for identifying and cloning peptide synthetase genes.  
 A:Reference number: S60747; MUID:95152187  
 A:Accession: S60747  
 A:Status: preliminary; nucleic acid sequence not shown  
 A:Molecule type: DNA  
 A:Residues: 1-142 <TUR>  
 A:Cross-references: EMBL:X78470  
 C:Superfamily: acetate-CoA ligase homology; acyl carrier protein homology  
 C:Keywords: carrier protein  
 F:1-102/Domain: acetate-CoA ligase homology (fragment) <ACL>  
 F:119-142/Domain: acyl carrier protein homology (fragment) <ACP>

Query Match 52.2%; Score 36; DB 2; Length 142;  
 Best Local Similarity 54.5%; Pred. No. 21;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 WYXXRGIRPYG 11  
 1:1111  
 Db 129 WEEVLGVRPYG 139

RESULT 15

T35841  
 Probable membrane protein - Streptomyces coelicolor  
 C:Species: Streptomyces coelicolor  
 C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
 C:Accession: T35841  
 R:Harris, D.; Taylor, K.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
 Submitted to the EMBL Data Library, August 1998  
 A:Reference number: Z21590  
 A:Accession: T35841  
 A:Status: preliminary; translated from GB/EMBL/DBJ  
 A:Molecule type: DNA  
 A:Residues: 1184 <HAR>  
 A:Cross-references: EMBL:AL031260; PIDN:CAA20292.1; GSPDB:GNO0070; SCOEDB:SC9A10.05C  
 A:Experimental source: strain A3(2)  
 C:Genetics:  
 A:Gene: SCOEDB:SC9A10.05C

Query Match 52.2%; Score 36; DB 2; Length 184;  
 Best Local Similarity 46.2%; Pred. No. 27;  
 Matches 6; Conservative 1; Mismatches 6; Indels 0; Gaps 0;  
 OY 1 WYXXRGIRPYG 13  
 1:1111  
 Db 91 WWSAHGVQVGF 103

Search completed: September 13, 2002, 09:24:01  
 Job time: 776 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:47 ; Search time 80.21 Seconds  
(without alignments)  
6.758 Million cell updates/sec

Title: US-09-446-543a-73-COPY\_8\_21  
Perfect score: 69  
Sequence: 1 WYXXRGIRPVGRFX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	66	95.7	87	PRRP_HUMAN
2	65	94.2	83	PRRP_RAT
3	65	94.2	98	PRRP_BOVIN
4	42	60.9	719	NRPI_YEAST
5	39	56.5	485	SYE_BACD
6	39	56.5	676	EXLI_HUMAN
7	38	55.1	637	MTVA_DROME
8	38	55.1	2832	NDVR_RHIME
9	37	53.6	149	ENRN_BPT7
10	37	53.6	788	REC2_HABIN
11	36	52.2	402	EX7L_STRCO
12	35	50.7	179	RK6_GUTR
13	35	50.7	391	THIL_ZOORA
14	35	50.7	453	TBB2_GEOCN
15	35	50.7	460	SR54_HALNI
16	35	50.7	618	CIRI_CITFR
17	35	50.7	922	GYRA_AERSA
18	35	50.7	1187	TYK2_HUMAN
19	34	49.3	224	YGCI_ECOLI
20	34	49.3	306	RSL_SYNP6
21	34	49.3	328	RSIA_SYNY3
22	34	49.3	424	SAMB_SALTY
23	34	49.3	801	TFR2_HUMAN
24	34	49.3	943	SYI_PSEFL
25	34	49.3	987	EPB4_MOUSE
26	34	49.3	1046	HHUA_HABIN
27	34	49.3	1077	HGRA_HABIN
28	33	47.8	254	EMD_HUMAN
29	33	47.8	270	CH14_PHAVU
30	33	47.8	424	PIP_AEROS
31	33	47.8	443	ARGA_ECOLI
32	33	47.8	477	PPOX_HUMAN
33	33	47.8	477	PPOX_MOUSE

34	33	47.8	582	1	KICH_YEAST	P20485	saccharomyc
35	33	47.8	636	1	ENV_MCFE	P15073	mlnk cell f
36	33	47.8	781	1	PBPB_HABIN	P45345	p penicillia
37	33	47.8	850	1	PBPA_ECOLI	P02918	escherichia
38	33	47.8	972	1	HGBA_HABDU	O47952	haemophilus
39	33	47.8	972	1	HGBB_HABDU	O47957	haemophilus
40	33	47.8	987	1	EPB4_HUMAN	P54760	homo sapien
41	33	47.8	1723	1	AIMI_HUMAN	O944X1	p genome po
42	32.5	47.1	3061	1	POLG_PYYHU	O02963	p genome po
43	32	46.4	150	1	ULAT_HCMVA	P16828	human cytom
44	32	46.4	177	1	NUSG_CAMJE	O9P136	campylobact
45	32	46.4	249	1	Y361_SYNY3	O35378	synechocyst

## ALIGNMENTS

RESULT	ID	PRRP_HUMAN	STANDARD	PRT	87 AA.
AC	P81277				
DT	30-MAY-2000	(Rel. 39, Created)			
DT	30-MAY-2000	(Rel. 39, Last sequence update)			
DT	01-MAR-2002	(Rel. 41, Last annotation update)			
DE	Prolactin-releasing peptide precursor (PrRP)	(Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP1; Prolactin-releasing peptide PrRP20].			
GN	PRH.				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.				
OX	NCBI_Taxid:9606;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RA	MEDLINE=98266781; PubMed=9607765;				
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;				
RA	"A prolactin-releasing peptide in the brain.";				
RT	Nature 393:272-276(1998).				
RL	[2]				
RN	TISSUE SPECIFICITY.				
RP	PubMed=10498338;				
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,				
RA	Sumino Y., Fujino M.;				
RT	"Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor.";				
RL	Regul. Pept. 83:1-10(1999).				
CC	-1- FUNCTION: Stimulates prolactin (PrL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PrL.				
CC	-1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a> or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> ).				
CC	-----				
DR	EMBL: AB015419; BAA29027.1;				
DR	MM: 602663;				
KW	Hormone; Amidation; Signal.				
FT	SIGNAL				
FT	PEPTIDE				
FT	PEPTIDE				
FT	MOD_RES				
SQ	SEQUENCE				

BY SIMILARITY.  
PROLACTIN-RELEASING PEPTIDE PRRP1.  
PROLACTIN-RELEASING PEPTIDE PRRP20.  
AMIDATION (G-54 PROVIDE AMIDE GROUP).

Query Match 95.7%; Score 66; DB 1; Length 87;  
 Best Local Similarity 84.6%; Pred. No. 3e-06;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
 II |IIIIIIII  
 DB 41 WYASRGIRPVGRF 53

## RESULT 2

PRRP\_RAT STANDARD: PRT: 83 AA.  
 ID PRRP\_RAT  
 AC P81278:  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-releasing peptide PRRP20].  
 DE PRH.  
 GN Rattus norvegicus (Rat).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsunoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 RN [2]  
 RP TISSUE SPECIFICITY.  
 RX PubMed=10498338;  
 RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RT "Tissue distribution of prolactin-releasing peptide (PRP) and its receptor.";  
 RL Regul. Pept. 83:1-10(1999).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 -----

EMBL: AB015418; BAA29026.1; -  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 FT SIGNAL 1 21 BY SIMILARITY.  
 FT PEPTIDE 22 52 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE 33 52 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 83 AA; 9215 MW; DDC75A264EEB4F29 CRC64;

Query Match 94.2%; Score 65; DB 1; Length 83;  
 Best Local Similarity 84.6%; Pred. No. 4.5e-06;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
 II |IIIIIIII  
 DB 40 WYTGIRPVGRF 52

RESULT 3  
 PRRP\_BOVIN STANDARD: PRT: 98 AA.  
 ID PRRP\_BOVIN  
 AC P81264:  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-releasing peptide PRRP20].  
 DE PRH.  
 GN Bos taurus (Bovine).  
 OS Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OX Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsunoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@sdb-sdb.ch](mailto:license@sdb-sdb.ch)).  
 -----

EMBL: AB015417; BAA29025.1; -  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA508 CRC64;

Query Match 94.2%; Score 65; DB 1; Length 98;  
 Best Local Similarity 84.6%; Pred. No. 5.3e-06;  
 Matches 11; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGRF 13  
 II |IIIIIIII  
 DB 41 WYASRGIRPVGRF 53

## RESULT 4

NRPI\_YEAST STANDARD: PRT: 719 AA.  
 ID NRPI\_YEAST  
 AC P32770; Q12228;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Asparagine-rich protein (ARP protein).  
 GN NRPI OR ARP OR ARP OR YDL167C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
 OX NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=AH22;

```

RX MEDLINE-93247548; PubMed-8483449;
RA Wehner E.P., Rao E., Brendel M.;
RT "Molecular structure and genetic regulation of SPA, a gene
RT responsible for resistance to formaldehyde in Saccharomyces
RT cerevisiae, and characterization of its protein product.";
RL Mol. Gen. Genet. 237:351-358(1993).
RN
RP SEQUENCE FROM N.A.
RC STRAIN-S288C;
RA Pohl T.M.;
RL Submitted (NOV-1995) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: CONTAINS 2 RANBP2-TYPE ZINC FINGERS.
CC -1- SIMILARITY: CONTAINS 1 RNA RECOGNITION MOTIF (RRM).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X68020: CAA48159.1; -
DR EMBL: Z67750: CAA91579.1; -
DR EMBL: Z74215: CAA98741.1; -
DR PIR: S31139; S31139.
DR HSSP: P04170; GRXN.
DR SCD: S0002326; NRPL.
DR InterPro: IPR000504; RRM.
DR InterPro: IPR001876; Znf-RanBP.
DR Pfam: PF00076; rrm; 1.
DR Pfam: PF00641; zf-RanBP; 2.
DR SMART: SM00360; RRM; 1.
DR SMART: SM00547; znf_RBZ; 2.
DR PROSITE: PS50102; RRM; 1.
DR PROSITE: PS00030; RRM_RNP_1; FALSE_NEG.
DR PROSITE: PS01358; zf_RANBP2_1; 2.
DR PROSITE: PS50199; zf_RANBP2_2; 2.
DR NCBI: ProStar; znc-finger; RNA-binding; Repeat.
KW Nucleic protein; zinc-finger; RNA-binding; Repeat.
FT DOMAIN 226 322 RNA-BINDING (RRM).
FT ZN_FING 355 384 RANBP2-TYPE 1.
FT ZN_FING 581 610 RANBP2-TYPE 2.
FT DOMAIN 490 564 ASN-RICH.
FT CONFLICT 493 493 I -> N (IN REF. 1).
SQ SEQUENCE 719 AA; 79299 MW; ADA9BC09F5D582669 CRC64;

Query Match 60.98; Score 42; DB 1; Length 719;
Best Local Similarity 54.58; Pred. No. 1.5;
Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVG 11
   1: 1:|||||
Db 245 WFTQYGVPRVG 255

RESULT 5
ID SYE_BACHD STANDARD; PRT; 485 AA.
AC 09KGF6;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA Ligase)
DE (GLURS).
GN GTFX OR BH0109.
OS Bacillus halodurans.
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Bacillus.
OC Bacillus/Staphylococcus group; Bacillus.
OX NCBI_TaxID=86665;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-C-125 / JCM 9153;

```

```

RX MEDLINE-20512582; PubMed-11058132;
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
RA Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,
RA Horikoshi K.;
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
RT halodurans and genomic sequence comparison with Bacillus subtilis.";
RL Nucleic Acids Res. 28:4317-4331(2000).
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +
CC diphosphate + L-glutamyl-tRNA(Glu).
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AP001507: BAB03828.1; -
DR InterPro: IPR000924; tRNA-synt_1c.
DR InterPro: IPR001412; tRNA-synt_1.
DR Pfam: PF00749; tRNA-synt_1c; 1.
DR PRINTS: PR00987; TRNASYNTGLU.
DR PROSITE: PS00178; AA-TRNA-LIGASE_I; 1.
KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;
KW Complete proteome.
FT SITE 11 "HIGH" REGION.
FT SITE 252 "KMSK" REGION.
FT BINDING 255 256 ATP (BY SIMILARITY).
SQ SEQUENCE 485 AA; 54785 MW; 7D34A86218F5786 CRC64;

```

```

Query Match 56.58; Score 39; DB 1; Length 485;
Best Local Similarity 58.38; Pred. No. 3.9;
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 2 YXXRGIRPVGRF 13
   1: 1:|||||
Db 146 YEAKGIRPVGRF 157

RESULT 6
ID EX1L_HUMAN STANDARD; PRT; 676 AA.
AC Q92935;
DT 15-JUL-1998 (Rel. 36, Created)
DT 15-JUL-1998 (Rel. 36, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Exostosin-like 1 (Exostosin-L) (Multiple exostosin-like protein).
GN EXTL1 OR EXT1.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
AC MEDLINE-97189339; PubMed-9037597;
RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;
RT "Identification and localization of the gene for EXTL1, a third member
RT of the multiple exostosins gene family.";
RL Genome Res. 7:10-16(1997).
RN
RP SEQUENCE FROM N.A.
RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;
RT "Mutations of the EXT genes in hereditary multiple exostoses in
RT Chinese.";
RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Wuyts W., Spieker N., Van Roy N., De Paep A., De Boulle K.,

```

RA Willems P.J., Van Hul W., Versteeg R., Speleman F.:  
 RT Refined physical mapping and genomic structure of the EXT1 gene."  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67191; AAC51141.1; -.  
 DR EMBL: AF083623; AAD02840.1; -.  
 DR EMBL: AF083623; AAD02840.1; JOINED.  
 DR EMBL: AF083624; AAD02840.1; JOINED.  
 DR EMBL: AF083625; AAD02840.1; JOINED.  
 DR EMBL: AF083626; AAD02840.1; JOINED.  
 DR EMBL: AF083627; AAD02840.1; JOINED.  
 DR EMBL: AF083628; AAD02840.1; JOINED.  
 DR EMBL: AF083629; AAD02840.1; JOINED.  
 DR EMBL: AF083630; AAD02840.1; JOINED.  
 DR EMBL: AF083631; AAD02840.1; JOINED.  
 DR EMBL: AF083632; AAD02840.1; JOINED.  
 DR EMBL: AF153980; AAF73172.1; -.  
 DR EMBL: AF153981; AAF73172.1; JOINED.  
 DR MIM: 601738; -.  
 DR InterPro: IPR004263; Exostosin.  
 DR Pfam: PF03016; Exostosin; 1.  
 KW Anti-oncogene; Multigene family; Transmembrane; Signal-anchor.  
 FT TRANSMEM 10 30  
 FT SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT CAROYND 269 269 N-LINKED (GLCNAC...) (POTENTIAL).  
 FT SEQUENCE 676 AA; 74673 MW; B5E006A8762E3633 CRC64;  
 SQ  
 Query Match 56.5%; Score 39; DB 1; Length 676;  
 Best local similarity 53.8%; Pred. No. 5.5;  
 Matches 7; Conservative 2; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 WYXXKGRIPVGRF 13  
 DB 407 YVLOQGSREGRF 419  
 RESULT 7  
 MTHA\_DROME STANDARD; PRT; 637 AA.  
 ID MTHA\_DROME  
 AC Q9W0R5;  
 DT 01-MAR-2002 (rel. 41, Created)  
 DT 01-MAR-2002 (rel. 41, Last sequence update)  
 DT 01-MAR-2002 (rel. 41, Last annotation update)  
 DE Probable G-protein-coupled receptor Mth-like 10 precursor (Methuselah-  
 DE like 10 protein).  
 GN MTHL10 OR CG17061.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Anthropoda; Tracheata; Hexapoda; Insecta;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 CC Ephydroidea; Drosophilidae; Drosophila.  
 OK NCBI\_TaxId=7227;  
 RN [1]  
 RC SEQUENCE FROM N.A.  
 RP STRAIN=Berkeley;  
 RX MEDLINE-20196006; PubMed-10731132;  
 RA Adams M.D., Geinlker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scher S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,

RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abrial J.F., Aghayani A., An H.-J., Andrews-Planckoch C., Baldwin D.,  
 RA Bailey R.M., Basu A., Baxendale J., Bayraktaroglu K., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borzova D., Botchan M.R., Bouck J., Brokstein P., Brotter P.,  
 RA Burtis K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahke C., Davenport L.B., Davies S.M.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kenison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kuip D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Modarity C., Morris J., Moshireli A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Relvert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster."  
 RL Science 287:2185-2195(2000).  
 RN [2]  
 RP REVIEW.  
 RA MEDLINE-20370890; PubMed-10908591;  
 RA Brody T., Cravchik A.;  
 RT "Drosophila melanogaster G protein-coupled receptors."  
 RL J. Cell Biol. 150:F83-F88(2000).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: BELONGS TO FAMILY 2 OF G-PROTEIN COUPLED  
 CC RECEPTORS. WITH SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AE003468; AAF47378.1; -.  
 DR FlyBase: FBgn0035132; CG17061.  
 DR InterPro: IPR000832; GPCR\_secretin.  
 DR Pfam: PF00002; 7tm\_2; 1.  
 DR PROSITE: PS00649; G\_PROTEIN\_REC\_F2\_1; FALSE\_NEG.  
 DR PROSITE: PS00650; G\_PROTEIN\_REC\_F2\_2; FALSE\_NEG.  
 DR PROSITE: PS00651; G\_PROTEIN\_REC\_F2\_4; 1.  
 KW Receptor; G-protein coupled receptor; Transmembrane; Glycoprotein;  
 KW Signal; Multigene family.  
 FT SIGNAL 1 32  
 FT CHAIN 33 637  
 FT FT  
 FT DOMAIN 33 235  
 FT TRANSMEM 236 256  
 FT DOMAIN 257 267  
 FT TRANSMEM 268 288  
 FT DOMAIN 289 299  
 FT TRANSMEM 300 320  
 FT DOMAIN 321 340  
 FT TRANSMEM 341 361  
 FT DOMAIN 362 391  
 FT TRANSMEM 392 412  
 FT  
 POTENTIAL.  
 PROBABLE G-PROTEIN-COUPLED RECEPTOR MTH-  
 LIKE 10.  
 EXTRACELLULAR (POTENTIAL).  
 1 (POTENTIAL).  
 CYTOPLASMIC (POTENTIAL).  
 2 (POTENTIAL).  
 EXTRACELLULAR (POTENTIAL).  
 3 (POTENTIAL).  
 CYTOPLASMIC (POTENTIAL).  
 4 (POTENTIAL).  
 EXTRACELLULAR (POTENTIAL).  
 5 (POTENTIAL).

FT DOMAIN 413 443 CYTOPLASMIC (POTENTIAL).  
FT TRANSSEM 444 464 6 (POTENTIAL).  
FT DOMAIN 465 475 EXTRACELLULAR (POTENTIAL).  
FT TRANSSEM 476 496 7 (POTENTIAL).  
FT DOMAIN 497 637 CYTOPLASMIC (POTENTIAL).  
FT DISULFID 56 110 BY SIMILARITY.  
FT DISULFID 112 117 BY SIMILARITY.  
FT DISULFID 121 216 BY SIMILARITY.  
FT DISULFID 122 135 BY SIMILARITY.  
FT DISULFID 177 236 BY SIMILARITY.  
FT CARBOHYD 63 63 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 72 72 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 142 142 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 152 152 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 157 157 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 198 198 N-LINKED (GLCNAC. . .) (POTENTIAL).  
FT CARBOHYD 223 223 N-LINKED (GLCNAC. . .) (POTENTIAL).  
SQ SEQUENCE 637 AA; 72988 MW; C55FA399B15CE777 CRC64;

Query Match 55.1%; Score 38; DB 1; Length 637;  
Best Local Similarity 46.2%; Pred. No. 8.2;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Oy 1 WYXXRGIRPVGRF 13  
Db 323 WHNFRGTGRGINRF 335

RESULT 8  
NDVB\_RHIME STANDARD; PRT; 2832 AA.  
ID NDVB\_RHIME  
AC P20471;  
DT 01-FEB-1991 (Rel. 17, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Protein ndvb.  
GN NDVB OR R03286 OR SMC04382.  
OS Rhizobium meliloti (Sinorhizobium meliloti).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_Taxid=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90153914; Pubmed=2154461;  
RA Ielpi L., Dylan T., Ditta G.S., Heliński D.R., Stanfield S.W.;  
RT "The ndvb locus of Rhizobium meliloti encodes a 319-kDa protein  
involved in the production of beta-(1-->2)-glucan.";  
RL J. Biol. Chem. 265:2843-2851(1990).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX STRAIN=1021;  
RC MEDLINE=21396507; Pubmed=11481430;  
RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
RA Boistard P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S.,  
RA Godtse T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masuy D.,  
RA Pohl T., Portetle D., Puehler A., Purnelle B., Ramepeter U.,  
RA Renard C., Thebault P., Vandenbol M., Weidner S., Gilbert F.;  
RT "Analysis of the chromosome sequence of the legume symbiont  
Sinorhizobium meliloti strain 1021".  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
CC -1- FUNCTION: INVOLVED IN THE PRODUCTION OF BETA-(1,2)-GLUCAN.  
CC IT IS INVOLVED NOT ONLY IN INVASION BUT ALSO IN BACTEROID  
CC DEVELOPMENT.  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
CC -1- SIMILARITY: TO A.TUMERICIENS CHVB.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL: J05219; AAA26305.1; ALT\_INIT.  
DR EMBL: AL591793; CAC47865.1; -.  
DR PIR: A35548; A35548.  
KW Transmembrane; Inner membrane; Complete proteome.  
FT TRANSSEM 411 431 POTENTIAL.  
FT TRANSSEM 444 464 POTENTIAL.  
FT TRANSSEM 810 830 POTENTIAL.  
FT TRANSSEM 831 851 POTENTIAL.  
FT TRANSSEM 959 979 POTENTIAL.  
FT CONFLICT 47 47 T -> A (IN REF. 1).  
FT CONFLICT 1187 1187 L -> F (IN REF. 1).  
FT CONFLICT 2602 2602 D -> G (IN REF. 1).  
SQ SEQUENCE 2832 AA; 315746 MW; 171EA89F03A936F2 CRC64;

Query Match 55.1%; Score 38; DB 1; Length 2832;  
Best Local Similarity 50.0%; Pred. No. 38;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 WYXXRGIRPV 10  
Db 1088 WYTRGLEPM 1097

RESULT 9  
ENRN\_BPT7 STANDARD; PRT; 149 AA.  
ID ENRN\_BPT7  
AC P00641;  
DT 21-JUL-1986 (Rel. 01, Created)  
DT 21-JUL-1986 (Rel. 01, Last sequence update)  
DT 01-NOV-1995 (Rel. 32, Last annotation update)  
DE Endonuclease I (EC 3.1.21.2) (Endonuclease).  
GN 3.  
OS Bacteriophage T7.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like phages.  
OX NCBI\_Taxid=10760;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=83241725; Pubmed=6864790;  
RA Dunn J.J., Studier F.W.;  
RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
locations of T7 genetic elements.";  
RL J. Mol. Biol. 166:477-535(1983).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=82078034; Pubmed=7310871;  
RA Dunn J.J., Studier F.W.;  
RT "Nucleotide sequence from the genetic left end of bacteriophage T7  
DNA to the beginning of gene 4.";  
RL J. Mol. Biol. 148:303-330(1981).  
CC -1- FUNCTION: ENDONUCLEASE I, WHICH IS EXPRESSED IN THE LATE  
CC STAGE, IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN  
CC OF HOST DNA. IN THE EARLY STAGE OF INFECTION, T7 DNA REPLICATES AS  
CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA  
CC LINEAR CONCATAMERS SEVERAL GENOMES IN LENGTH. THE GENE 3 PRODUCT  
CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE CONCATAMERS.  
CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
CC phosphooligonucleotide end-products.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>)  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
CC -----  
DR EMBL: V01127; CAA24345.1; -;  
DR EMBL: V01146; CAA24402.1; -;  
DR PIR: A00785; NEBP37.

DR PIR: S42301; S42301.  
 KW Hydrolase; Nuclease; Endonuclease.  
 SQ SEQUENCE 149 AA; 17172 MW; D092AA28E3743BC1 CRC64;

Query Match 53.6%; Score 37; DB 1; Length 149;  
 Best Local Similarity 58.3%; Pred. No. 2.9;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 YXXRGIRPGVF 13  
 | :||| | | |  
 DB 4 YGAKGIRKVCAP 15

RESULT 10  
 REC2\_HAEIN STANDARD; PRT: 788 AA.  
 AC P4408;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-NOV-1995 (Rel. 32, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Recombination protein 2.  
 GN REC2 OR REC-2 OR HI0061.  
 OS Haemophilus influenzae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
 OC Haemophilus.  
 OX NCBI\_TaxID=727;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RC200;  
 RX MEDLINE-94341577; PubMed-8063112;  
 RA Clifton S.W., McCarthy D., Roe B.A.;  
 RT "Sequence of the rec-2 locus of Haemophilus influenzae: homologues to  
 come-ORF3 of Bacillus subtilis and msba of Escherichia coli.";  
 RL Gene 146:95-100(1994).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-RD / KW20 / ATCC 51907;  
 RX MEDLINE-95350630; PubMed-7542800;  
 RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
 RA Kierleavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
 RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
 RA Scott J.D., Shirley R., Liu L.-T., Glodek A., Kelley J.M.,  
 RA Weidman J.F., Phillips C.A., Spriggs T., Heddlom E., Cotton M.D.,  
 RA Utechtack T.R., Hanna M.C., Nguyen D.T., Saudex D.M., Brandon R.C.,  
 RA Fine L.D., Fritchman J.L., Fuhman J.L., Geoghegan N.S.M.,  
 RA Gnehm J.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
 RA Venter J.C.;  
 RT "Whole-genome random sequencing and assembly of Haemophilus  
 influenzae Rd.";  
 RL Science 269:496-512(1995).  
 CC -1- FUNCTION: MIGHT CONTRIBUTE TO TRANSFORMATION AS A MEMBER OF A  
 MEMBRANE BOUND PORE COMPLEX AT THE BASE OF THE TRANSFORMASOME. IT  
 COULD DIRECTLY INTERACT WITH TRANSFORMING DNA DURING TRANSLOCATION  
 INDIRECTLY BY PARTICIPATING IN THE ASSEMBLY OF THE PORE.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
 (Potential).  
 CC -1- SIMILARITY: TO B. SUBTILIS COMEC, N. GONORRHORAE COMA, AND E. COLI  
 YCAL.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: L20805; AAC13733.1; -;  
 DR EMBL: 032691; AAC21739.1; -;  
 DR TIGR: HI0061; -;  
 DR InterPro: IPR001279; Beta\_lactam\_mct.  
 DR Pfam: PF00753; lactamase\_B; 1.

KW Transport; Transmembrane; Inner membrane; Complete proteome.  
 FT TRANSMEM 3 23 POTENTIAL.  
 FT TRANSMEM 27 47 POTENTIAL.  
 FT TRANSMEM 51 71 POTENTIAL.  
 FT TRANSMEM 226 246 POTENTIAL.  
 FT TRANSMEM 263 283 POTENTIAL.  
 FT TRANSMEM 313 333 POTENTIAL.  
 FT TRANSMEM 367 387 POTENTIAL.  
 FT TRANSMEM 401 421 POTENTIAL.  
 FT TRANSMEM 434 454 POTENTIAL.  
 FT TRANSMEM 456 476 POTENTIAL.  
 FT TRANSMEM 508 528 POTENTIAL.  
 FT CONFLICT 748 788  
 FT VENTAVSGVRYNFEQRLLEIQARTKESPMYARVIGLSKE  
 FT -> GRKYRCFASAGKFFSRPIRNPASSHRIPLVCACNMI  
 FT IKIGKTMRAIFIR (IN REF. 1).  
 SQ SEQUENCE 788 AA; 89355 MW; F31104595CB4E47A CRC64;

Query Match 53.6%; Score 37; DB 1; Length 788;  
 Best Local Similarity 54.5%; Pred. No. 16;  
 Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 WYXXRGIRPGV 11  
 | :||| | | |  
 DB 156 WYFSKGITAVG 166

RESULT 11  
 EX7L\_STRCO STANDARD; PRT: 402 AA.  
 ID EX7L\_STRCO  
 AC Q9FBM3;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable exodeoxyribonuclease VII large subunit (EC 3.1.11.6)  
 DE (Exonuclease VII large subunit).  
 GN XSEA OR SCR7.29c.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D., Cerdano A.M., Parthill J., Barrell B.G.,  
 RA Rajadream M.A.;  
 RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: BIDIRECTIONALLY DEGRADES SINGLE-STRANDED DNA INTO LARGE  
 ACID-INSOLUBLE OLIGONUCLEOTIDES, WHICH ARE THEN DEGRADED FURTHER  
 INTO SMALL ACID-SOLUBLE OLIGONUCLEOTIDES (BY SIMILARITY).  
 CC -1- CATALYTIC ACTIVITY: Exonucleolytic cleavage in either 5' to 3'-  
 or 3' to 5'-direction to yield 5'-phosphomononucleotides.  
 CC -1- SUBUNIT: HETEROOLIGOMER COMPOSED OF LARGE AND SMALL SUBUNITS (BY  
 SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE XSEA FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC EMBL: AL391754; GAC05901.1; -;  
 DR InterPro: IPR003753; Exonuc\_VII\_L.  
 DR InterPro: IPR002309; tRNA-synt\_2.  
 DR Pfam: PF02601; Exonuc\_VII\_L.  
 DR Pfam: PF01336; tRNA\_anti\_1.  
 KW Hydrolase; Nuclease; Exonuclease.  
 SQ SEQUENCE 402 AA; 43882 MW; 145929A8372B4E08 CRC64;

Query Match 52.2%; Score 36; DB 1; Length 402;  
Best Local Similarity 42.1%; Pred. No. 13;  
Matches 8; Conservative 1; Mismatches 2; Indels 8; Gaps 1;

QY 1 WYXR-----IRPVG 11  
DB 91 WYAPRGQLSLRAAEIRKPVG 109

## RESULT 12

RC\_GUITH 12  
ID RK6\_GUITH STANDARD; PRT; 179 AA.  
AC 046908;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 15-JUL-1999 (Rel. 38, Last annotation update)  
DE Chloroplast 50S ribosomal protein L6.  
GN RPL6.  
OS Guillardia theta (Cryptomonas phl).  
OC Chloroplast.  
OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
OX NCBI\_TaxID=55529;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97283757; PubMed=9137835;  
RA Wang S.L., Liu X.-O., Douglas S.E.;  
RT "The large ribosomal protein gene cluster of a cryptomonad plastid:  
RT gene organization, sequence and evolutionary implications.";  
RL Blochem. Mol. Biol. Int. 41:1035-1044(1997).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99128221; PubMed=9929392;  
RA Douglas S.E., Penny S.L.;  
RT "The plastid genome of the cryptophyte alga, Guillardia theta:  
RT complete sequence and conserved synteny groups confirm its common  
RT ancestry with red algae.";  
RL J. Mol. Evol. 48:236-244(1999).  
CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
CC LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE  
CC PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF041468; AAC35717.1; -.  
DR HSSP: P02391; 1RL6.  
DR InterPro: IPR000702; Ribosomal\_L6.  
DR InterPro: IPR002358; Ribosomal\_L6\_1.  
DR Pfam: PF00347; Ribosomal\_L6; 1.  
DR PRINTS: PR00059; RIBOSOMAL\_L6.  
DR PRODOM: PD002236; RIBOSOMAL\_L6; 1.  
DR PROSITE: PS00525; RIBOSOMAL\_L6\_1; 1.  
KM Ribosomal protein: Chloroplast.  
SQ SEQUENCE 179 AA; 19527 MW; 8B4C0DBD0152AD24 CRC64;

Query Match 50.7%; Score 35; DB 1; Length 179;  
Best Local Similarity 58.3%; Pred. No. 8.8;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

QY 2 YXRGIRPVGRF 13  
DB 157 YKGGIRYVGEF 168

## RESULT 13

THIL\_ZOORA  
ID THIL\_ZOORA STANDARD; PRT; 391 AA.  
AC P07097;  
DT 01-APR-1988 (Rel. 07, Created)  
DT 01-APR-1990 (Rel. 14, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE Acetyl-CoA acetyltransferase (EC 2.3.1.9) (Acetoacetyl-CoA thiolase).  
GN PHBA.  
OS Zoogloea ramifera.  
OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;  
OC Zoogloea.  
OX NCBI\_TaxID=350;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ATCC 19623 / I-16-M;  
RX MEDLINE=87083504; PubMed=2878929;  
RA Peoples O.P., Masamune S., Walsh C.T., Sinskey A.J.;  
RT "Biosynthetic thiolase from Zoogloea ramifera. III. Isolation and  
RT characterization of the structural gene.";  
RL J. Biol. Chem. 262:97-102(1987).  
RN [2]  
RP REVISION TO 130.  
RX MEDLINE=89359356; PubMed=2670935;  
RA Peoples O.P., Sinskey A.J.;  
RT "Poly-beta-hydroxybutyrate biosynthesis in Alcaligenes eutrophus H16.  
RT Characterization of the genes encoding beta-ketothiolase and  
RT acetoacetyl-CoA reductase.";  
RL J. Biol. Chem. 264:15293-15297(1989).  
RN [3]  
RP MUTAGENESIS OF CYS-377.  
RX MEDLINE=91217075; PubMed=1673680;  
RA Palmer M.A.J., Differding E., Gamboni R., Williams S.F., Peoples O.P.,  
RA Walsh C.T., Sinskey S.J., Masamune S.;  
RT "Biosynthetic thiolase from Zoogloea ramifera. Evidence for a  
RT mechanism involving Cys-378 as the active site base.";  
RL J. Biol. Chem. 266:8369-8375(1991).  
CC -1- CATALYTIC ACTIVITY: 2 acetyl-CoA + CoA + acetoacetyl-CoA.  
CC -1- PATHWAY: FIRST STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.  
CC -1- SUBUNIT: HOMOTETRAMER.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: J02631; AAA27706.1; ALT\_SEQ.  
DR PIR: A26121; XGXZAC.  
DR PIR: A27754; A27754.  
DR HSSP: P27796; 1AFT.  
DR InterPro: IPR002155; Thiolase.  
DR Pfam: PF00108; thiolase; 1.  
DR Pfam: PF02803; thiolase\_C; 1.  
DR PROSITE: PS00098; THIOLASE\_1; 1.  
DR PROSITE: PS00099; THIOLASE\_3; 1.  
DR PROSITE: PS00737; THIOLASE\_2; 1.  
KW transferase; Acyltransferase; PHB biosynthesis.  
FT INIT\_MET 0  
FT ACT\_SITE 88  
FT ACT\_SITE 377  
FT ACT\_SITE 377  
FT MUTAGEN 377  
SQ SEQUENCE 391 AA; 40342 MW; 6D2351A1BC0E4EDD CRC64;

Query Match 50.7%; Score 35; DB 1; Length 391;  
Best Local Similarity 75.0%; Pred. No. 20;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 5 RGIRPVGR 12

```

Db      266 KGIQPLGR 273

      ||::||
      ||::||
RESULT  14
TBB2_GEOCN STANDARD: PRT: 453 AA.
AC P32925:
DT 01-OCT-1993 (Rel. 27, Created)
DT 01-NOV-1995 (Rel. 32, Last sequence update)
DT 15-JUL-1999 (Rel. 38, Last annotation update)
DE Tubulin beta-2 chain.
OS Geotrichum candidum (Oospora laetis).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Dipodascaceae; Galactomycetes.
OX NCBI_TaxID=27317;
RN      [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=92079883; PubMed=1836049;
RA Gold S.E., Casale W.L., Keen N.T.;
RT Characterization of two beta-tubulin genes from Geotrichum
   candidum.
RL Mol. Genet. 230:104-112(1991).
CC -1- FUNCTION: TUBULIN IS THE MAJOR CONSTITUENT OF MICROTUBULES. IT
CC BINDS TWO MOLES OF GTP, ONE AT AN EXCHANGEABLE SITE ON THE BETA
CC CHAIN AND ONE AT A NONEXCHANGEABLE SITE ON THE ALPHA-CHAIN.
CC -1- SUBUNIT: DIMER OF ALPHA AND BETA CHAINS.
CC -1- SIMILARITY: BELONGS TO THE TUBULIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: S69627; AAB20557.1; ALT_SEQ.
DR PIR: S18597; S18597.
DR InterPro: IPR002453; Beta_tubulin.
DR InterPro: IPR000217; Tubulin.
DR InterPro: IPR003008; Tubulin_Ftsz.
DR Pfam: PF00091; tubulin.1.
DR PRINTS: PR01161; TUBULIN.
DR PROSITE: PS00227; TUBULIN.1.
DR PROSITE: PS00228; TUBULIN_B_AUTOREG.1.
DR Microtubules; GTP-binding; Multigene family.
DR NP_BIND 142 148 GTP (POTENTIAL).
FT SEQUENCE 453 AA: 50399 MW: 2EA9D8A0246E0371 CRC64;
SQ
Query Match 50.7%; Score 35; DB 1; Length 453;
Best Local Similarity 50.0%; Pred. No. 23;
Matches 5; Conservative 1; Mismatches 4; Indels 0; Gaps 0;
QY 1 WYXXRGIRPV 10
   ||::||
   ||::||
Db      399 WYTGEMEPV 408

RESULT  15
SR54_HALNT STANDARD: PRT: 460 AA.
AC Q9HNM5:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Signal recognition 54 kDa protein (SRP54).
DE SRP54 OR VNG2459G.
OS Halobacterium sp. (strain NRC-1).
OC Archaea; Euryarchaeota; Halobacteriales; Halobacteriaceae;
OC Halobacterium.
OX NCBI_TaxID=64091;

```

```

RN      [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20504483; PubMed=11016950;
RA Ng W.V., Kennedy S.P., Maharis G.G., Bergquist B., Pan M.,
RA Shukla H.D., Lasky S.R., Haliga N.S., Thorsson V., Sbrogna J.,
RA Swartzell S., Weir D., Hall J., Dahl T.A., Weir R., Goo Y.A.,
RA Leithausen B., Keller K., Cruz R., Danon M.J., Hough D.W.,
RA Maddocks D.G., Jablonski P.E., Krebs M.P., Angevine C.M., Dale H.,
RA Isenbarger T.A., Peck R.F., Pohlschroder M., Spudis J.L., Jung K.-H.,
RA Alam M., Freitas T., Hou S., Daniels C.J., Dennis P.P., Omer A.D.,
RA Ehardt H., Lowe T.M., Liang P., Riley M., Hood L., Dassarma S.;
RT "Genome sequence of Halobacterium species NRC-1."
RL Proc. Natl. Acad. Sci. U.S.A. 97:12176-12181(2000).
CC -1- FUNCTION: BINDS TO THE SIGNAL SEQUENCE OF PRESECRETORY PROTEIN
CC WHEN THEY EMERGE FROM THE RIBOSOMES (BY SIMILARITY).
CC -1- SUBUNIT: ARCHAEL SIGNAL RECOGNITION PARTICLE CONSISTS OF A 7S RNA
CC MOLECULE OF 300 NUCLEOTIDES AND TWO PROTEIN SUBUNITS: SRP54 AND
CC SRP19 (BY SIMILARITY).
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
CC -1- DOMAIN: HAS A TWO DOMAIN STRUCTURE: THE G-DOMAIN BINDS GTP; THE
CC M-DOMAIN BINDS THE 7S RNA IN PRESENCE OF SRP19 AND ALSO BINDS THE
CC SIGNAL SEQUENCE (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE SRP FAMILY OF GTP-BINDING PROTEINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation-
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@isb-sdb.ch).
CC -----
DR EMBL: AE005125; AAG20536.1; -.
DR InterPro: IPR000897; SRP54.
DR Pfam: PF00448; SRP54.1.
DR ProDom: PD000819; SRP54.1.
DR PROSITE: PS00300; SRP54; FALSE_NEG.
DR Signal recognition particle; GTP-binding; RNA-binding;
KW Complete proteome.
DR DOMAIN 1 289
FT DOMAIN 290 460 M-DOMAIN (BY SIMILARITY).
FT NP_BIND 104 111 GTP (BY SIMILARITY).
FT NP_BIND 184 188 GTP (BY SIMILARITY).
FT NP_BIND 242 245 GTP (BY SIMILARITY).
FT DOMAIN 449 458 POLY-GLY.
SQ SEQUENCE 460 AA: 50149 MW: 8361B782E651352E CRC64;
QY 1 WYXXRGIRPV 9
   ||::||
   ||::||
Db      120 WFSKGLRP 128

Query Match 50.7%; Score 35; DB 1; Length 460;
Best Local Similarity 44.4%; Pred. No. 23;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

```

Search completed: September 13, 2002, 09:30:47  
Job time: 1137 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:21 ; Search time 311.85 Seconds  
(without alignments)  
7.766 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_8\_21  
Perfect score: 69  
Sequence: 1 WYXXRGIRPVGRFX 14

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacterioplasmid:\*  
17: sp\_archaeoplasmid:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	62	89.9	117	13	09W624
2	40	62.3	420	2	0931Z7
3	40	58.0	153	9	09T133
4	39	56.5	256	16	09ZVE9
5	38	55.1	333	16	09PH76
6	38	55.1	481	5	09SWT7
7	38	55.1	536	5	09SYN4
8	38	55.1	536	5	09SPK7
9	38	55.1	536	5	09SPK6
10	38	55.1	536	5	09SPK5
11	38	55.1	536	5	09SPK7
12	38	55.1	536	5	09SPK6
13	38	55.1	536	5	09SPK5
14	38	55.1	536	5	09SPK7
15	37	53.6	332	5	09W043
16	37	53.6	338	16	09ZXB6

17	37	53.6	343	16	09ZXB4	09ZXB4 rhizobium m
18	37	53.6	348	5	09W1G4	09W1G4 drosophila
19	37	53.6	357	5	09S0U7	09S0U7 drosophila
20	37	53.6	395	2	054095	054095 streptomyces
21	37	53.6	769	2	068843	068843 streptomyces
22	37	53.6	885	16	09ZXB4	09ZXB4 rhizobium m
23	37	53.6	1450	11	054728	054728 rat mus norv
24	36.5	52.9	163	2	046610	046610 desulfovibrio
25	36	52.2	184	2	045285	045285 bacillus 11
26	36	52.2	184	2	086838	086838 streptomyces
27	36	52.2	275	16	053929	053929 mycobacterium
28	36	52.2	293	16	09PHR2	09PHR2 campylobacter
29	36	52.2	300	16	09ZSA6	09ZSA6 rhizobium m
30	36	52.2	302	16	09ZSA6	09ZSA6 rhizobium m
31	36	52.2	338	5	018729	018729 caenorhabditis
32	36	52.2	399	2	093EC5	093EC5 rhizobium 1
33	36	52.2	426	17	0978S0	0978S0 thermoplasma
34	36	52.2	480	12	09PYX8	09PYX8 xestia c-nl
35	36	52.2	488	16	0997D8	0997D8 staphylococcus
36	36	52.2	503	16	09KRY1	09KRY1 vibrio chol
37	36	52.2	540	10	09LGD0	09LGD0 oryza sativ
38	36	52.2	545	16	09A7W7	09A7W7 caulobacter
39	36	52.2	641	2	093CZ5	093CZ5 streptomyces
40	36	52.2	674	16	050431	050431 mycobacterium
41	36	52.2	790	10	09W371	09W371 arabidopsis
42	36	52.2	925	5	09VXK4	09VXK4 drosophila
43	36	52.2	954	5	097180	097180 drosophila
44	36	52.2	966	2	047673	047673 escherichia
45	36	52.2	975	13	P79750	P79750 fuqu rufip

#### ALIGNMENTS

RESULT 1  
ID 09W624 PRELIMINARY; PRT; 117 AA.  
AC 09W624;  
DT 01-NOV-1999 (TREMblrel. 12, Created)  
DT 01-NOV-1999 (TREMblrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
DE C-RE AMIDE.  
OS Carassius auratus (goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
CC Actinopterygii; Cyprinodontiformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=BRAIN;  
RA Satoh H., Minakata H., Fujimoto M.;  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AB020024; BAA6662.1; -  
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 89.9%; Score 62; DB 13; Length 117;  
Best Local Similarity 69.2%; Pred. No. 0.00026;  
Matches 9; Conservative 2; Mismatches 2; Indels 0; Gaps 0;  
QY 1 WYXXRGIRPVGRFX 13  
DB 63 WYXGVRPGRFX 75  
RESULT 2  
ID 0931Z7 PRELIMINARY; PRT; 420 AA.  
AC 0931Z7;  
DT 01-DEC-2001 (TREMblrel. 19, Created)  
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria: Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CCM339;  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CCM339.";  
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF033994; AAK61719.1; -  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 62.3%; Score 43; DB 2; Length 420;  
Best Local Similarity 100.0%; Pred. No. 4.8;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 6 GIRPVGRF 13  
| | | | | | | |  
DB 48 GIRPVGRF 55

RESULT 3  
O9T133 PRELIMINARY; PRT; 153 AA.  
AC O9T133;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DR 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
GN ENDONUCLEASE.  
OS Bacteriophage phlyeo3-12.  
OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
OC T7-like phages.  
OX NCBI\_TaxID=110457;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Pajunen M.I.;  
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21125554; PubMed=11222590;  
RA Pajunen M.I., Kiljunen S.J., Soederholm M.E.L., Skurnik M.;  
RT "Complete genomic sequence of the lytic bacteriophage phlyeo3-12 of  
RT Yersinia enterocolitica serotype O:3.";  
RL J. Bacteriol. 183:1928-1937(2001).  
DR EMBL; AJ251805; CAB63604.1; -  
SQ SEQUENCE 153 AA; 17640 MW; 211571BBDE6C641D CRC64;

Query Match 58.0%; Score 40; DB 9; Length 153;  
Best Local Similarity 58.3%; Pred. No. 6;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 2 YXXRGIRPYGRF 13  
| | | | | | | |  
DB 5 YXRGYRKYGAF 16

RESULT 4  
O92VE9 PRELIMINARY; PRT; 256 AA.  
ID O92VE9;  
AC O92VE9;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DR 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
GN HYPOTHETICAL PROTEIN SM21253.  
RT Rhizobium meliloti (Sinorhizobium meliloti).

OG Plasmid pSymB (megaplasmid 2).  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Sinorhizobium.  
OX NCBI\_TaxID=382;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=1021;  
RX MEDLINE=21396508; PubMed=11481431;  
RA Finan T.M., Weldner S., Wong K., Buhrmester J., Chain P.,  
RA Vorholter F.J., Hernandez-Lucas I., Becker A., Cowle A., Gouzy J.,  
RA Golding B., Puchler A.;  
RT "The complete sequence of the 1,683-kb pSymB megaplasmid from the N2-  
RT fixing endosymbiont Sinorhizobium meliloti.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:9889-9894(2001).  
DR EMBL; AL603644; CAC49157.1; -  
KW Plasmid; Hypothetical protein; Complete proteome.  
SQ SEQUENCE 256 AA; 28459 MW; 54B064834CEC7C39 CRC64;

Query Match 56.5%; Score 39; DB 16; Length 256;  
Best Local Similarity 54.5%; Pred. No. 16;  
Matches 6; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 WYXXRGIRPYG 11  
| : | | | | | |  
DB 185 WHGTGRCRPYG 195

RESULT 5  
O9PH76 PRELIMINARY; PRT; 333 AA.  
ID O9PH76;  
AC O9PH76;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DR 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
GN HYDROXYBENZONATE OCTAPRENTLTRANSFERASE.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,  
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincani A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franco S.C., Franco M.C., Frohme M., Furlan L.R.,  
RA Gartner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohnselt J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madela A.M.B.N., Madela H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nham J.A., Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A., Jr., Pezquerro J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E., Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silva J.C.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuchiko M.H.,  
RA Vallada H., Van Sluys M.A., Verjovsky-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
RL Nature 406:151-159(2000).

DR EMBL: AE003860; AAF82881.1; -  
DR InterPro: IPR000537; UBLA.  
DR Pfam: PF01040; UBLA; 1.  
DR PROSITE: PS00943; UBLA; UNKNOWN\_1.  
DR Complete proteome.  
SQ SEQUENCE 333 AA; 37931 MW; ECF3F4716C962B95 CRC64;

Query Match 55.1%; Score 38; DB 16; Length 333;  
Best Local Similarity 63.6%; Pred. No. 34;  
Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVG 11  
DB 58 WKLAGDRPVG 68

RESULT 6  
Q95WT7 ID Q95WT7 PRELIMINARY; PRT; 481 AA.  
AC Q95WT7;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR PROTEIN (FRAGMENT).  
GN MTH.  
OS Drosophila yakuba (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7245;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YAK\_30;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF300415; AAK97895.1; -  
KW Receptor.  
FT NON\_TER  
SQ SEQUENCE 481 AA; 55732 MW; 44773A7D5203F3FB CRC64;

Query Match 55.1%; Score 38; DB 5; Length 481;  
Best Local Similarity 46.2%; Pred. No. 51;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
DB 297 WHNFRGTRGINRF 309

RESULT 7  
Q95YNA ID Q95YNA PRELIMINARY; PRT; 536 AA.  
AC Q95YNA;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR.  
GN MTH.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=ZIM(S)24;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF300396; AAK97885.1; -  
KW Receptor.  
SQ SEQUENCE 536 AA; 61782 MW; FDE6174267361561 CRC64;

KW Receptor.  
SQ SEQUENCE 536 AA; 61803 MW; E74D7E65B4F2E1P1 CRC64;

Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
DB 297 WHNFRGTRGINRF 309

RESULT 8  
Q95PK7 ID Q95PK7 PRELIMINARY; PRT; 536 AA.  
AC Q95PK7;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR PROTEIN.  
GN MTH.  
OS Drosophila yakuba (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7245;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YAK\_4;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF300417; AAK97896.1; -  
KW Receptor.  
SQ SEQUENCE 536 AA; 61806 MW; 335CD38CB919F734 CRC64;

Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

QY 1 WYXXRGIRPVGRF 13  
DB 297 WHNFRGTRGINRF 309

RESULT 9  
Q95PK6 ID Q95PK6 PRELIMINARY; PRT; 536 AA.  
AC Q95PK6;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR PROTEIN.  
GN MTH.  
OS Drosophila yakuba (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7245;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YAK\_5;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog.";  
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF300419; AAK97897.1; -  
KW Receptor.  
SQ SEQUENCE 536 AA; 61782 MW; FDE6174267361561 CRC64;

Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGR 13  
I: ||| : ||  
Db 297 WHNFRGTRGIRNF 309

RESULT 10  
Q95PK5 PRELIMINARY; PRT; 536 AA.

AC Q95PK5  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR PROTEIN.  
GN MTH.  
OS Drosophila yakuba (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7245;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YAK\_8;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog."  
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF300421; AAK97898.1; -  
DR EMBL; AF300420; AAK97898.1; JOINED.  
KW Receptor.  
SQ SEQUENCE 536 AA; 61795 MW; 58B6086E67361567 CRC64;

Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGR 13  
I: ||| : ||  
Db 297 WHNFRGTRGIRNF 309

RESULT 11  
Q95NU7 PRELIMINARY; PRT; 536 AA.

AC Q95NU7  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VARIOUS STRAINS;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog."  
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF300407; AAK97891.1; -  
DR EMBL; AF300364; AAK97869.1; -  
DR EMBL; AF300363; AAK97869.1; JOINED.  
DR EMBL; AF300366; AAK97870.1; -  
DR EMBL; AF300365; AAK97870.1; JOINED.  
DR EMBL; AF300368; AAK97871.1; -  
DR EMBL; AF300367; AAK97871.1; JOINED.  
DR EMBL; AF300370; AAK97872.1; -

DR EMBL; AF300369; AAK97872.1; JOINED.  
DR EMBL; AF300372; AAK97873.1; -  
DR EMBL; AF300371; AAK97873.1; JOINED.  
DR EMBL; AF300374; AAK97874.1; -  
DR EMBL; AF300373; AAK97874.1; JOINED.  
DR EMBL; AF300376; AAK97875.1; -  
DR EMBL; AF300375; AAK97875.1; JOINED.  
DR EMBL; AF300378; AAK97876.1; -  
DR EMBL; AF300377; AAK97876.1; JOINED.  
DR EMBL; AF300380; AAK97877.1; -  
DR EMBL; AF300382; AAK97877.1; JOINED.  
DR EMBL; AF300382; AAK97878.1; -  
DR EMBL; AF300381; AAK97878.1; JOINED.  
DR EMBL; AF300384; AAK97879.1; -  
DR EMBL; AF300383; AAK97879.1; JOINED.  
DR EMBL; AF300386; AAK97880.1; -  
DR EMBL; AF300385; AAK97880.1; JOINED.  
DR EMBL; AF300388; AAK97881.1; -  
DR EMBL; AF300387; AAK97881.1; JOINED.  
DR EMBL; AF300390; AAK97882.1; -  
DR EMBL; AF300389; AAK97882.1; JOINED.  
DR EMBL; AF300392; AAK97883.1; -  
DR EMBL; AF300391; AAK97883.1; JOINED.  
DR EMBL; AF300394; AAK97884.1; -  
DR EMBL; AF300393; AAK97884.1; JOINED.  
DR EMBL; AF300398; AAK97886.1; -  
DR EMBL; AF300397; AAK97886.1; JOINED.  
DR EMBL; AF300400; AAK97887.1; -  
DR EMBL; AF300399; AAK97887.1; JOINED.  
DR EMBL; AF300402; AAK97888.1; -  
DR EMBL; AF300401; AAK97888.1; JOINED.  
DR EMBL; AF300404; AAK97889.1; -  
DR EMBL; AF300403; AAK97889.1; JOINED.  
DR EMBL; AF300406; AAK97890.1; -  
DR EMBL; AF300405; AAK97890.1; JOINED.  
KW Receptor.  
SQ SEQUENCE 536 AA; 61835 MW; 4C4A03D0410C105C CRC64;

Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

OY 1 WYXXRGIRPVGR 13  
I: ||| : ||  
Db 297 WHNFRGTRGIRNF 309

RESULT 12  
Q95NT6 PRELIMINARY; PRT; 536 AA.

AC Q95NT6  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DE G PROTEIN-COUPLED RECEPTOR PROTEIN.  
GN MTH.  
OS Drosophila yakuba (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxId=7245;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=YAK\_23, BG103, AND YAK\_2;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog."  
RL Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AF300413; AAK97894.1; -  
DR EMBL; AF300413; AAK97894.1; JOINED.  
DR EMBL; AF300410; AAK97892.1; -  
DR EMBL; AF300409; AAK97892.1; JOINED.  
DR EMBL; AF300412; AAK97893.1; -

DR EMBL; AF300411; AAK97893.1; JOINED.  
KW RECEPTOR.  
SQ SEQUENCE 536 AA; 61768 MW; 58E6124767361567 CRC64;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 WYXXRGIRPVGRF 13  
Db 297 WHNFRGTGRINRF 309  
RESULT 13  
ID 095NR7 PRELIMINARY; PRT; 536 AA.  
AC 095NR7;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE G-PROTEIN-COUPLED RECEPTOR PROTEIN.  
OS Drosophila simulans (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_Taxid=7240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C96.6.1S, HFL97.4.1S, GA96.2.1S, AND VA96.8.1S;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF300360; AAK82820.1; JOINED.  
DR EMBL; AF300342; AAK82811.1; JOINED.  
DR EMBL; AF300341; AAK82811.1; JOINED.  
DR EMBL; AF300344; AAK82812.1; JOINED.  
DR EMBL; AF300343; AAK82812.1; JOINED.  
DR EMBL; AF300354; AAK82817.1; JOINED.  
DR EMBL; AF300353; AAK82817.1; JOINED.  
KW Receptor.  
SQ SEQUENCE 536 AA; 61805 MW; 6F0F81C940995AEF CRC64;  
Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 WYXXRGIRPVGRF 13  
Db 297 WHNFRGTGRINRF 309  
RESULT 14  
ID 095N00 PRELIMINARY; PRT; 536 AA.  
AC 095N00;  
DT 01-DEC-2001 (TREMBlrel. 19, Created)  
DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE G-PROTEIN-COUPLED RECEPTOR PROTEIN.  
OS Drosophila simulans (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_Taxid=7240;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=VARIOUS STRAINS;  
RA Duvernell D.D., Schmidt P.S., Eanes W.F.;  
RT "Molecular population genetics of a methuselah paralog."  
RL Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AF300362; AAK82821.1; JOINED.  
DR EMBL; AF300361; AAK82821.1; JOINED.  
DR EMBL; AF300338; AAK82809.1; JOINED.  
DR EMBL; AF300337; AAK82809.1; JOINED.  
DR EMBL; AF300340; AAK82810.1; JOINED.  
DR EMBL; AF300339; AAK82810.1; JOINED.  
DR EMBL; AF300346; AAK82813.1; JOINED.  
DR EMBL; AF300345; AAK82813.1; JOINED.  
DR EMBL; AF300348; AAK82814.1; JOINED.  
DR EMBL; AF300347; AAK82814.1; JOINED.  
DR EMBL; AF300350; AAK82815.1; JOINED.  
DR EMBL; AF300349; AAK82815.1; JOINED.  
DR EMBL; AF300352; AAK82816.1; JOINED.  
DR EMBL; AF300351; AAK82816.1; JOINED.  
DR EMBL; AF300356; AAK82818.1; JOINED.  
DR EMBL; AF300355; AAK82818.1; JOINED.  
DR EMBL; AF300358; AAK82819.1; JOINED.  
DR EMBL; AF300357; AAK82819.1; JOINED.  
KW Receptor.  
SQ SEQUENCE 536 AA; 61839 MW; 86A1852837B132E1 CRC64;  
Query Match 55.1%; Score 38; DB 5; Length 536;  
Best Local Similarity 46.2%; Pred. No. 58;  
Matches 6; Conservative 2; Mismatches 5; Indels 0; Gaps 0;  
OY 1 WYXXRGIRPVGRF 13  
Db 297 WHNFRGTGRINRF 309  
RESULT 15  
ID 09W043 PRELIMINARY; PRT; 332 AA.  
AC 09W043;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CG5707 PROTEIN.  
GN BEST:HL03644 OR CG5707.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriere S., Fleischmann W.,  
RA Foster C., Gabriellian A.E., Garcia N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gottrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Honick J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jaisel M., Kalush F., Karen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Klumel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Lio X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,

RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spler E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svitskaas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissendbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003474; AAF47616.1; -  
DR FlyBase: FBgn0026593; BEST:HL03644.  
DR InterPro: IPR000073; Abhydrolase.  
DR InterPro: IPR000379; Est\_Llp\_thloest\_actate.  
DR Pfam: PF00561; abhydrolase; 1.  
SQ FDCI56A4D36E56F CRC64;  
SEQUENCE 332 AA; 38665 MW; FDCI56A4D36E56F CRC64;

Query Match 53.6%; Score 37; DB 5; Length 332;  
Best Local Similarity 50.0%; Pred No. 53;  
Matches 5; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

Oy 1 WYXXRGIRPV 10  
|||:  
Db 10 WYGNRQVRPI 19

Search completed: September 13, 2002, 09:29:22  
Job time: 1067 sec



PT studying diseases related to ligand abnormality  
XX  
XX  
PS Disclosure: Page 26; 73pp; Japanese.  
CC  
XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAV49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA:  
  
Query Match 94.8%; Score 55; DB 21; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00035;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
DB 4 yasrgirpvgrf 15  
  
RESULT 2  
AAV49296  
ID AAV49296 standard; peptide: 15 AA.  
XX  
AC AAV49296;  
XX  
XX 22-FEB-2000 (first entry)  
XX  
DE 19p2 ligand peptide fragment.  
XX  
XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
XX pituitary; regulatory mechanism; central nervous system; pancreatic.  
OS Homo sapiens.  
XX  
XX  
FH Key Location/Qualifiers  
FT Modified-site 15 /note="C-terminal amide"  
FT  
XX  
XX WO9960112-A1.  
XX  
XX 25-NOV-1999.  
XX  
XX 20-MAY-1999; 99WO-JP02650.  
XX  
XX 21-MAY-1998; 98JP-0140293.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
XX  
XX WPI: 2000-039381/03.  
XX  
XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
XX studying diseases related to ligand abnormality  
XX  
XX  
XX Disclosure: Page 27; 73pp; Japanese.  
XX  
XX The invention provides a monoclonal antibody which has a specific  
XX reaction with the part peptide of the C-terminal of 19p2 ligand or its  
XX derivative. The antibodies can be used in diagnosis or to treat or  
XX prevent diseases associated with abnormality in the pituitary function  
XX regulatory mechanism (e.g. promotion of prolactin secretion), central  
XX nervous regulatory mechanism, and pancreatic function regulatory  
XX mechanism. The antibody-based immunoassay can also be applied in  
XX clarifying the physiological functions of the ligand and its derivative.  
XX Sequences AAV49290-302 represent peptide fragments of the 19p2 ligand.

XX  
SQ Sequence 15 AA:  
  
Query Match 94.8%; Score 55; DB 21; Length 15;  
Best Local Similarity 83.3%; Pred. No. 0.00035;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
DB 4 yasrgirpvgrf 15  
  
RESULT 3  
AAW31394  
ID AAW31394 standard; Peptide: 20 AA.  
XX  
AC AAW31394;  
XX  
XX 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
XX modulator; pituitary; central nervous system; pancreas; prophylactic;  
XX therapeutic agent.  
XX  
XX Homo sapiens.  
XX  
XX  
XX WO9724436-A2.  
XX  
XX 10-JUL-1997.  
XX  
XX 26-DEC-1996; 96WO-JP03821.  
XX  
XX 18-SEP-1996; 96JP-0246573.  
XX  
XX 28-DEC-1995; 95JP-0345371.  
XX  
XX 15-MAR-1996; 96JP-0059419.  
XX  
XX 12-AUG-1996; 96JP-0211805.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
XX Kawamata Y, Kitada C;  
XX  
XX WPI: 1997-363672/33.  
XX  
XX N-PSDB; AAV02431.  
XX  
XX Ligand peptide for G protein-coupled receptor - acts by modulating  
XX function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
XX Claim 2; Page 185; 258pp; English.  
XX  
XX This sequence represents a peptide fragment from a novel human type  
XX ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
XX sequence represented in AAW31390 and is used in an assay to monitor  
XX ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
XX compositions containing this ligand may be used as a pituitary function  
XX modulator, a central nervous system modulator or a pancreatic function  
XX modulator. This ligand could have specific applications as a  
XX prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
XX syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
XX trauma, growth hormone secretory disease, hyper- and polyphagia,  
XX hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
XX hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
XX Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
XX transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
XX acute myocardial infarction, infertility, spinocerebellar degeneration,  
XX bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
XX oligodactylia. Assays can also be developed to screen compounds which are  
XX capable of altering the binding activity of the ligand affecting  
XX activation of the G protein-coupled receptor protein.



SQ Sequence 20 AA:

Query Match 94.8%: Score 55; DB 18; Length 20;  
Best Local Similarity 83.3%: Pred. No. 0.00048;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
I | | | | | | | | | |  
Db 9 yasrgirpvgrf 20

RESULT 4  
AAW97236  
ID AAW97236 standard; peptide; 20 AA.  
XX  
AC AAW97236;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Human type ligand polypeptide fragment.  
XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Charli-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome or dysospermia;  
KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.  
XX  
OS Homo sapiens.  
XX  
PN W09858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX  
DR WPI; 1999-105614/09.  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 166; 241pp; English.  
XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Charli-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dysospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, interruption mole, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX

SQ Sequence 20 AA:

Query Match 94.8%: Score 55; DB 20; Length 20;  
Best Local Similarity 83.3%: Pred. No. 0.00048;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
I | | | | | | | | | |  
Db 9 yasrgirpvgrf 20

RESULT 5  
AAB10365  
ID AAB10365 standard; peptide; 20 AA.  
XX  
AC AAB10365;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 35.  
XX  
KW Human; oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN W0200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PF 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98UP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Matsumoto H, Kitada C, Hinuma S;  
XX  
DR WPI; 2000-452298/39.  
XX  
PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
XX  
PS Disclosure; Page 63; 72pp; Japanese.  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
SQ Sequence 20 AA:

Query Match 94.8%: Score 55; DB 21; Length 20;  
Best Local Similarity 83.3%: Pred. No. 0.00048;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
I | | | | | | | | | |  
Db 9 yasrgirpvgrf 20

```
RESULT 6
AAV49294 standard; peptide: 20 AA.
XX AAV49294;
AC AAV49294;
DT 22-FEB-2000 (first entry)
DE 19p2 ligand peptide fragment.
XX
XX Monoclonal antibody: 19p2 ligand; diagnosis: prolactin secretion;
KW pituitary; regulatory mechanism: central nervous system; pancreatic.
XX
XX Homo sapiens.
OS
FH Key Location/Qualifiers
FT Modified-site 20 /note="C-terminal amide"
XX
XX W09960112-A1.
PN
XX 25-NOV-1999.
PD
XX
XX 20-MAY-1999: 99WO-JP02650.
XX
XX 21-MAY-1998: 98JP-0140293.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Matsumoto H, Kitada C, Hinuma S;
PI WPI: 2000-039381/03.
XX
XX New monoclonal antibodies, useful in diagnosis, as drugs and in
PT studying diseases related to ligand abnormality -
XX
XX Disclosure: Page 26; 73pp; Japanese.
PS
XX
XX The invention provides a monoclonal antibody which has a specific
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its
CC derivative. The antibodies can be used in diagnosis or to treat or
CC prevent diseases associated with abnormality in the pituitary function
CC regulatory mechanism (e.g. promotion of prolactin secretion), central
CC nervous regulatory mechanism, and pancreatic function regulatory
CC mechanism. The antibody-based immunoassay can also be applied in
CC clarifying the physiological functions of the ligand and its derivative.
CC Sequences AAV49290-302 represent peptide fragments of the 19p2 ligand.
XX
XX
XX Sequence 20 AA:

Query Match 94.8%; Score 55; DB 21; Length 20;
Best Local Similarity 83.3%; Pred. No. 0.00048;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXRGIRPVGRF 12
   | | | | | | | |
DB 9 yasrgirpvgrf 20

RESULT 7
AAG62534 standard; peptide: 20 AA.
XX AAG62534;
AC AAG62534;
DT 24-AUG-2001 (first entry)
DE Human CRH releasing protein related peptide SEQ ID NO: 35.
XX
XX Human: corticotrophin releasing hormone; CRH: G protein receptor ligand;
KW analgesic; hyperaldosteronism; hypercortisolaemia; hypoadrenocorticism;
KW Addison's disease; adrenal gland hyperfunction; obesity.
XX
```

```
XX
XX Homo sapiens.
OS
XX W0200135984-A1.
PN
XX 25-MAY-2001.
PD
XX
XX 17-NOV-2000; 2000WO-JP08119.
XX
XX 18-NOV-1999: 99JP-0327900.
XX
XX 26-SEP-2000; 2000JP-0297073.
XX
XX (TAKE ) TAKEDA CHEM IND LTD.
XX
XX Kitada C, Matsumoto H, Hinuma S;
PI WPI: 2001-355552/37.
XX
XX Use of G protein receptor ligand or peptide for controlling
PT corticotrophin releasing hormone secretion -
XX
XX Claim 4; Page 75; 90pp; Japanese.
PS
XX
XX The present sequence describes a method of controlling the secretion of
CC corticotrophin releasing hormone (CRH), involving the use of a G protein
CC receptor ligand. This can be used to control the secretion of CRH and is
CC useful as an analgesic or for treating, preventing or ameliorating
CC diseases associated with CRH secretion such as hyperaldosteronism,
CC hypercortisolaemia, secondary or chronic hypoadrenocorticism, Addison's
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair
CC loss, and hypotension), adrenal gland hypofunction and obesity. The
CC present sequence is a peptide used in the exemplification of the
CC invention.
XX
XX
XX Sequence 20 AA:

Query Match 94.8%; Score 55; DB 22; Length 20;
Best Local Similarity 83.3%; Pred. No. 0.00048;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXRGIRPVGRF 12
   | | | | | | | |
DB 9 yasrgirpvgrf 20

RESULT 8
AAB90992 standard; peptide: 20 AA.
XX AAB90992;
AC AAB90992;
DT 22-JUN-2001 (first entry)
DE Prolactin releasing peptide SEQ ID NO:166.
XX
XX Protection; endogenous therapeutic peptide; peptidase; conjugation;
KW blood component; modification: succinimidyl; maleimido group; amino;
KW hydroxyl; thiol; hormone; growth factor; neurotransmitter.
XX
XX Homo sapiens.
OS
XX Synthetic.
XX
XX W0200069900-A2.
PN
XX 23-NOV-2000.
PD
XX
XX 17-MAY-2000; 2000WO-US13576.
XX
XX 17-MAY-1999: 99US-0134406.
XX
XX 10-SEP-1998: 99US-0153406.
XX
XX 15-OCT-1999: 99US-0159783.
XX
```

PA (CONJ-) CONNUTCHEM INC.  
XX  
PI Bridon DP, Ezrin AM, Milner PG, Holmes DL, Thibaudau K;  
XX  
DR WPI; 2001-112059/12.  
XX  
PT Modifying and attaching therapeutic peptides to albumin prevents  
PT peptidase degradation, useful for increasing length of in vivo activity  
XX  
PS Disclosure; Page 244; 733pp; English.  
XX  
CC The present invention describes a modified therapeutic peptide (I)  
CC comprising a therapeutically active amino acid region (III) and a  
CC reactive group (II) (e.g. succinimidyl and maleimido groups) attached to  
CC a less therapeutically active amino acid region (IV), which covalently  
CC bonds with amino/hydroxyl/thiol groups on blood components to form a  
CC peptidease stabilised therapeutic peptide composed of 3-50 amino acids.  
CC (I) are useful for modifying therapeutic peptides e.g. hormones, growth  
CC factors and neurotransmitters, to protect them from peptidase activity  
CC in vivo for the treatment of various disorders. Endogenous therapeutic  
CC peptides are not suitable as drug candidates as they require frequent  
CC administration due to rapid degradation by peptidases in the body.  
CC Modifying and attaching therapeutic peptides to albumin prevents or  
CC reduces the action of peptidases to increase length of activity (half  
CC life) and specifically as bonding to large molecules decreases  
CC intercellular uptake and interference with physiological processes.  
CC AAB90829 to AAB92441 represent peptides which can be used in the  
CC exemplification of the present invention.  
XX  
SQ Sequence 20 AA:  
  
Query Match 94.8%; Score 55; DB 22; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00048;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 YXKRGIRPVGRF 12  
I | | | | | | | | | |  
Db 9 yasrglrpvgrf 20  
  
RESULT 9  
AAB31395  
ID AAB31395 standard; Peptide; 21 AA.  
XX  
AC AAB31395;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 5.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KM therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kltada C;

XX  
DR WPI; 1997-363672/33.  
DR N-PDB; AAV02432.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 54 of the  
CC sequence represented in AAB31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polynephria,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosacchara. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 21 AA:  
  
Query Match 94.8%; Score 55; DB 18; Length 21;  
Best Local Similarity 83.3%; Pred. No. 0.00051;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 YXKRGIRPVGRF 12  
I | | | | | | | | | |  
Db 9 yasrglrpvgrf 20  
  
RESULT 10  
AAB10366  
ID AAB10366 standard; peptide; 21 AA.  
XX  
AC AAB10366;  
XX  
DT 24-NOV-2000 (first entry)  
XX  
DE Human oxytocin secretion promoting peptide SEQ ID NO: 36.  
XX  
KM Human: oxytocin secretion promoter; G protein-coupled receptor protein;  
KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KM caesarean section; artificial fertilization; galactostasis; goat; pig;  
KM veterinary medicine; milk production.  
XX  
OS Homo sapiens.  
XX  
PN WO200038704-A1.  
XX  
PD 06-JUL-2000.  
XX  
PE 22-DEC-1999; 99WO-JP07199.  
XX  
PR 25-DEC-1998; 98JP-0369585.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
PI Matsumoto H, Kitada C, Hinuma S;  
PI WPI; 2000-452298/39.  
PT Physiologically-active polypeptide recognized as ligand by G

PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
XX medicine -  
XX  
PS Disclosure; Page 63; 72pp; Japanese.  
XX  
CC This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.  
XX  
SQ Sequence 21 AA:  
  
Query Match 94.8%; Score 55; DB 21; Length 21;  
Best Local Similarity 83.3%; Pred. No. 0.00051;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 YXXRGIRPYGRF 12  
I | | | | | | | | | |  
Db 9 yasrglrpygrf 20  
  
RESULT 11  
AAG62535  
ID AAG62535 standard; peptide: 21 AA.  
XX  
AC AAG62535;  
XX  
DT 24-AUG-2001 (first entry)  
XX  
DE Human CRH releasing protein related peptide SEQ ID NO: 36.  
XX  
XX Human; corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
XX Kawamata Y, Kitada C;  
XX WO200135984-A1.  
XX  
PD 25-MAY-2001.  
XX  
PF 17-NOV-2000; 2000MO-JP08119.  
XX  
XX 18-NOV-1999; 99JP-0327900.  
PR 26-SEP-2000; 2000JP-0297073.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Kitada C, Matsumoto H, Hinuma S;  
PI  
XX WPI; 2001-355552/37.  
DR  
XX  
XX Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
XX  
XX  
PS Disclosure; Page 75; 90pp; Japanese.  
XX  
CC The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH), involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including boredom, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The

CC present sequence is a peptide used in the exemplification of the  
CC invention.  
XX  
XX  
SQ Sequence 21 AA:  
  
Query Match 94.8%; Score 55; DB 22; Length 21;  
Best Local Similarity 83.3%; Pred. No. 0.00051;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 YXXRGIRPYGRF 12  
I | | | | | | | | | |  
Db 9 yasrglrpygrf 20  
  
RESULT 12  
AAW31396  
ID AAW31396 standard; peptide: 22 AA.  
XX  
AC AAW31396;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 6.  
XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
XX Homo sapiens.  
OS  
XX  
XX WO9724436-A2.  
XX  
XX 10-JUL-1997.  
XX  
PD 26-DEC-1996; 96MO-JP03821.  
XX  
XX 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
XX Kawamata Y, Kitada C;  
XX N-PSDB; AAW02433.  
XX  
DR WPI; 1997-363672/33.  
XX  
XX  
XX Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2; Page 186; 258pp; English.  
XX  
XX This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 55 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neuritis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligolactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting

CC activation of the G protein-coupled receptor protein.  
SQ Sequence 22 AA;

Query Match 94.8%; Score 55; DB 18; Length 22;  
Best Local Similarity 83.3%; Pred. No. 0.00053;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| |||||  
DB 9 yaargirpvgrf 20

RESULT 13  
AAB10367  
ID AAB10367 standard; peptide: 22 AA.

XX AAB10367;

XX 24-NOV-2000 (first entry)

DE Human oxytocin secretion promoting peptide SEQ ID NO: 37.

XX  
XX Human: oxytocin secretion promoter; G protein-coupled receptor protein;  
KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
KW veterinary medicine; milk production.

XX Homo sapiens.

OS  
PN WO200038704-A1.

XX 06-JUL-2000.

XX 22-DEC-1999; 99WO-JP07199.

XX 25-DEC-1998; 98JP-0369585.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-452298/39.

XX  
XX Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin.  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -

XX Disclosure; Page 64; 72pp; Japanese.

XX  
XX This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.,  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a human peptide which acts as an oxytocin secretion  
CC promoter.

XX Sequence 22 AA;

Query Match 94.8%; Score 55; DB 21; Length 22;  
Best Local Similarity 83.3%; Pred. No. 0.00053;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| |||||  
DB 9 yaargirpvgrf 20

RESULT 14  
AAG62536  
ID AAG62536 standard; peptide: 22 AA.

XX AAG62536;

XX 24-AUG-2001 (first entry)

DE Human CRH releasing protein related peptide SEQ ID NO: 37.

XX  
XX Human: corticotrophin releasing hormone; CRH; G protein receptor ligand;  
KW analgesic; hyperaldosteronism; hypercortisolemia; hypoadrenocorticism;  
KW Addison's disease; adrenal gland hyperfunction; obesity.

XX Homo sapiens.

OS  
PN WO200135984-A1.

XX 25-MAY-2001.

XX 17-NOV-2000; 2000WO-JP08119.

XX 18-NOV-1999; 99JP-0327900.

XX 26-SEP-2000; 2000JP-0297073.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Hinuma S;

XX WPI; 2001-355552/37.

XX  
XX Use of G protein receptor ligand or peptide for controlling  
PT corticotrophin releasing hormone secretion -  
PT  
XX Disclosure; Page 75; 90pp; Japanese.

XX  
XX The present sequence describes a method of controlling the secretion of  
CC corticotrophin releasing hormone (CRH). Involving the use of a G protein  
CC receptor ligand. This can be used to control the secretion of CRH and is  
CC useful as an analgesic or for treating, preventing or ameliorating  
CC diseases associated with CRH secretion such as hyperaldosteronism,  
CC hypercortisolemia, secondary or chronic hypoadrenocorticism, Addison's  
CC disease (including borden, nausea, pigmentation, hypogonadism, hair  
CC loss, and hypotension), adrenal gland hypofunction and obesity. The  
CC present sequence is a peptide used in the exemplification of the  
CC invention.

XX Sequence 22 AA;

Query Match 94.8%; Score 55; DB 22; Length 22;  
Best Local Similarity 83.3%; Pred. No. 0.00053;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| |||||  
DB 9 yaargirpvgrf 20

RESULT 15  
AAW31391  
ID AAW31391 standard; peptide: 31 AA.

XX AAW31391;

XX 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 1.

XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;

therapeutic agent.

OS Homo sapiens.  
 PN WO9724436-A2.  
 PD 10-JUL-1997.  
 PF 26-DEC-1996; 96WO-JP03821.  
 PR 18-SEP-1996; 96JP-0246573.  
 PR 28-DEC-1995; 95JP-0343371.  
 PR 15-MAR-1996; 96JP-0059419.  
 PR 12-AUG-1996; 96JP-0211805.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,  
 PI Kawamata Y, Kitada C;  
 DR WPI; 1997-363672/33.  
 DR N-PSDB; AAV02428.  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Claim 2; Page 184; 258pp; English.

This sequence represents a peptide fragment from a novel human type  
 CC ligand polypeptide corresponding to amino acid residues 23 to 53 of the  
 CC sequence represented in AAW1390 and is used in an assay to monitor  
 CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
 CC compositions containing this ligand may be used as a pituitary function  
 CC modulator, a central nervous system modulator or a pancreatic function  
 CC modulator. This ligand could have specific applications as a  
 CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
 CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
 CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
 CC hyperlipidemia, hypercholesterolemia, hyperglycemia,  
 CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
 CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
 CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
 CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
 CC oligosaccharia. Assays can also be developed to screen compounds which are  
 CC capable of altering the binding activity of the ligand affecting  
 CC activation of the G protein-coupled receptor protein.  
 XX  
 SQ Sequence 31 AA;

Query Match 94.8%; Score 55; DB 18; Length 31;  
 Best Local Similarity 83.3%; Pred. No. 0.00076;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
 I | | | | | | | | | |  
 DB 20 YASGIRPVGRF 31

Search completed: September 13, 2002, 09:18:35  
 Job time: 500 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 seconds  
(without alignments)  
2.324 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_9\_21

Perfect score: 58

Sequence: 1 YXXRGIRPYGRFX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /cgn2\_6/ptodata/2/1aa/5A\_COMB.pep.\*  
2: /cgn2\_6/ptodata/2/1aa/5B\_COMB.pep.\*  
3: /cgn2\_6/ptodata/2/1aa/6A\_COMB.pep.\*  
4: /cgn2\_6/ptodata/2/1aa/6B\_COMB.pep.\*  
5: /cgn2\_6/ptodata/2/1aa/PCMTUS\_COMB.pep.\*  
6: /cgn2\_6/ptodata/2/1aa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	55	94.8	20	3	US-09-105-678A-46
2	55	94.8	20	4	US-08-776-971-64
3	55	94.8	20	4	US-09-421-208-46
4	55	94.8	21	3	US-09-105-678A-47
5	55	94.8	21	4	US-08-776-971-65
6	55	94.8	21	4	US-09-421-208-47
7	55	94.8	22	3	US-09-105-678A-48
8	55	94.8	22	4	US-08-776-971-66
9	55	94.8	22	4	US-09-421-208-48
10	55	94.8	31	3	US-09-105-678A-9
11	55	94.8	31	3	US-09-105-678A-43
12	55	94.8	31	4	US-08-776-971-61
13	55	94.8	31	4	US-09-421-208-9
14	55	94.8	31	4	US-09-421-208-43
15	55	94.8	32	3	US-09-105-678A-44
16	55	94.8	32	4	US-08-776-971-62
17	55	94.8	32	4	US-09-421-208-44
18	55	94.8	33	3	US-09-105-678A-45
19	55	94.8	33	4	US-08-776-971-63
20	55	94.8	33	4	US-09-421-208-45
21	55	94.8	87	4	US-08-776-971-59
22	55	94.8	87	4	US-08-776-971-135
23	55	94.8	87	4	US-08-776-971-138
24	55	93.1	15	4	US-08-776-971-93
25	54	93.1	20	3	US-09-105-678A-34
26	54	93.1	20	3	US-09-105-678A-40
27	54	93.1	20	4	US-08-776-971-8

28	54	93.1	20	4	US-08-776-971-50	Sequence 50, Appl
29	54	93.1	20	4	US-08-776-971-98	Sequence 98, Appl
30	54	93.1	20	4	US-09-421-208-34	Sequence 34, Appl
31	54	93.1	20	4	US-09-421-208-40	Sequence 40, Appl
32	54	93.1	21	3	US-09-105-678A-35	Sequence 35, Appl
33	54	93.1	21	3	US-09-105-678A-41	Sequence 41, Appl
34	54	93.1	21	4	US-08-776-971-9	Sequence 9, Appl1
35	54	93.1	21	4	US-08-776-971-51	Sequence 51, Appl
36	54	93.1	21	4	US-09-421-208-35	Sequence 35, Appl
37	54	93.1	21	4	US-09-421-208-41	Sequence 41, Appl
38	54	93.1	22	3	US-09-105-678A-36	Sequence 36, Appl
39	54	93.1	22	3	US-09-105-678A-42	Sequence 42, Appl
40	54	93.1	22	4	US-08-776-971-10	Sequence 10, Appl
41	54	93.1	22	4	US-08-776-971-52	Sequence 52, Appl
42	54	93.1	22	4	US-09-421-208-36	Sequence 36, Appl
43	54	93.1	22	4	US-09-421-208-42	Sequence 42, Appl
44	54	93.1	31	3	US-09-105-678A-7	Sequence 7, Appl1
45	54	93.1	31	3	US-09-105-678A-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: Linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 94.8% Score 55; DB 3; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00017;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
DB 9 YASRGIRPVGRF 20

## RESULT 2

US-08-776-971-64  
; Sequence 64, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Kyo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776, 971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 64:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; FRAGMENT TYPE: Internal  
; SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 94.8%; Score 55; DB 4; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00017;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
DB 9 YASRGIRPVGRF 20

## RESULT 3

US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421, 208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105, 678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 46:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 94.8%; Score 55; DB 4; Length 20;  
Best Local Similarity 83.3%; Pred. No. 0.00017;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
DB 9 YASRGIRPVGRF 20

## RESULT 4

US-09-105-678A-47  
; Sequence 47, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston



```

: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/105,678A
: FILING DATE: 26-JUN-1998
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 47:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21 amino acids
: TYPE: amino acid
: STRANDEDNESS:
: TOPOLOGY: linear
: MOLECULE TYPE: peptide
: US-09-105-678A-47

```

```

Query Match          94.8%; Score 55; DB 3; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00018;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 YXXRGIRPVGRF 12
Db 9 YASRGIRPVGRF 20

```

```

RESULT 5
US-08-776-971-65
: Sequence 65, Application US/0876971B
: Patent No. 6228984
: GENERAL INFORMATION:
: APPLICANT: Hinuma, Shuji
:             Habata, Yugo
:             Kawamata, Yuji
:             Hosoya, Masaki
:             Fujii, Ryo
:             Fukusumi, Shoji
:             Kitada, Chieko
: TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE
: NUMBER OF SEQUENCES: 140
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Diskette
: COMPUTER: IBM compatible
: OPERATING SYSTEM: DOS
: SOFTWARE: FASTSEQ for Windows Version 2.0
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/776,971B
: FILING DATE: 06-Feb-1997
: CLASSIFICATION: <Unknown>
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: PCT/JP96/03821
: FILING DATE: 28-DEC-1996

```

```

: APPLICATION NUMBER: JP 7/343371
: FILING DATE: 28-DEC-1995
: APPLICATION NUMBER: JP 8/59419
: FILING DATE: 15-MAR-1996
: APPLICATION NUMBER: JP 8/211805
: FILING DATE: 12-AUG-1996
: APPLICATION NUMBER: JP 8/246573
: FILING DATE: 18-SEP-1996
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 47176
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400
: TELEFAX: 617-523-6440
: INFORMATION FOR SEQ ID NO: 65:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 21 amino acids
: TYPE: amino acid
: STRANDEDNESS: single
: TOPOLOGY: linear
: MOLECULE TYPE: protein
: FRAGMENT TYPE: internal
: SEQUENCE DESCRIPTION: SEQ ID NO: 65:
US-08-776-971-65

```

```

Query Match          94.8%; Score 55; DB 4; Length 21;
Best Local Similarity 83.3%; Pred. No. 0.00018;
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 YXXRGIRPVGRF 12
Db 9 YASRGIRPVGRF 20

```

```

RESULT 6
US-09-421-208-47
: Sequence 47, Application US/09421208
: Patent No. 6258561
: GENERAL INFORMATION:
: APPLICANT: Suenaga, Masato
:             Moriya, Takeo
:             Tanaka, Yoko
: APPLICANT: Nishimura, Osamu
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND
: NUMBER OF SEQUENCES: 52
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP
: STREET: 130 Water Street
: CITY: Boston
: STATE: MA
: COUNTRY: USA
: ZIP: 02109
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.30
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/09/421,208
: FILING DATE:
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 09/105,678
: FILING DATE: 26-JUN-1998
: APPLICATION NUMBER: JP 172118/1997
: FILING DATE: 27-JUN-1997
: ATTORNEY/AGENT INFORMATION:
: NAME: Conlin, David G.
: REGISTRATION NUMBER: 27,026
: REFERENCE/DOCKET NUMBER: 48466-342
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: 617-523-3400

```

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-47

Query Match 94.8%; Score 55; DB 4; Length 21;  
Best Local Similarity 83.3%; Pred. No. 0.00019;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
DB 9 YASRGIRPVGRF 20

RESULT 7  
US-09-105-678A-48  
Sequence 48, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Motiya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentia Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-48

Query Match 94.8%; Score 55; DB 3; Length 22;  
Best Local Similarity 83.3%; Pred. No. 0.00019;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
DB 9 YASRGIRPVGRF 20

RESULT 8  
US-08-776-971-66  
Sequence 66, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawabata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kikada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996

ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 66:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 66:  
US-08-776-971-66

Query Match 94.8%; Score 55; DB 4; Length 22;  
Best Local Similarity 83.3%; Pred. No. 0.00019;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
DB 9 YASRGIRPVGRF 20

RESULT 9  
US-09-421-208-48

Sequence 48, Application US/09421208  
Patent No. 6238561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 48:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 22 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-48

Query Match 94.8%; Score 55; DB 4; Length 22;  
Best Local Similarity 83.3%; Pred. No. 0.00019;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRCIRPVGRF 12  
| | | | | | | | | |  
DB 9 YASRCIRPVGRF 20

RESULT 10  
US-09-105-678A-9  
Sequence 9, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 9:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-9

Query Match 94.8%; Score 55; DB 3; Length 31;  
Best Local Similarity 83.3%; Pred. No. 0.00027;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 YXXRCIRPVGRF 12  
| | | | | | | | | |  
DB 20 YASRCIRPVGRF 31

RESULT 11  
US-09-105-678A-43  
Sequence 43, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 43:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678a-43

Query Match 94.8%; Score 55; DB 3; Length 31;  
Best Local Similarity 83.3%; Pred. No. 0.00027;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
Db 20 YASRGIRPVGRF 31

RESULT 12

US-08-776-971-61  
Sequence 61, Application US/08776971B  
Patent No. 6228984

GENERAL INFORMATION:

APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko

TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 61:

SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear

MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 61:  
US-08-776-971-61

Query Match 94.8%; Score 55; DB 4; Length 31;  
Best Local Similarity 83.3%; Pred. No. 0.00027;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
Db 20 YASRGIRPVGRF 31

RESULT 13

US-09-421-208-9  
Sequence 9, Application US/09421208  
Patent No. 6258561

GENERAL INFORMATION:

APPLICANT: Suenaga, Masato  
Moriya, Takeo  
Tanaka, Yoko

APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997

ATTORNEY/AGENT INFORMATION:

NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 9:

SEQUENCE CHARACTERISTICS:  
LENGTH: 31 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-9

Query Match 94.8%; Score 55; DB 4; Length 31;  
Best Local Similarity 83.3%; Pred. No. 0.00027;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| | | | | | | | | |  
Db 20 YASRGIRPVGRF 31

RESULT 14



**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:01; Search time 172.41 Seconds  
(without alignments)  
7.245 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_9\_21

Perfect score: 58

Sequence: 1 YXXRGIRPVGRFX 13

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Database: PIR\_71:\*  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	93.1	83	2 JC7607	prolactin-releasin
2	39	67.2	485	2 E83663	glutamyl-tRNA synt
3	37	63.8	149	1 NEBP37	endooxyribonucle
4	36	62.1	545	2 A87448	conserved hypothet
5	36	62.1	674	2 G70875	probable oxidoredu
6	36	62.1	790	2 T47959	hypothetical prote
7	35	60.3	178	2 D95026	ribosomal protein
8	35	60.3	178	2 D97897	50S ribosomal prot
9	35	60.3	272	2 T39533	hypothetical prote
10	35	60.3	391	1 XXGZAC	acetyl-CoA C-acety
11	35	60.3	504	2 B70821	hypothetical prote
12	35	60.3	519	2 T45447	probable two-compo
13	35	60.3	894	2 E82221	DNA gyrase, chain
14	35	60.3	1056	2 B82527	hypothetical prote
15	35	60.3	7463	2 T36248	CDA peptide synthe
16	34	58.6	226	2 A65057	hypothetical prote
17	34	58.6	230	2 F97646	hypothetical prote
18	34	58.6	290	2 AD2870	probable allphatic
19	34	58.6	307	2 S51485	ABC transporter, m
20	34	58.6	328	2 S77236	ribosomal protein
21	34	58.6	343	2 AH1823	30S ribosomal prot
22	33	56.9	68	2 D72428	hypothetical prote
23	33	56.9	313	2 B93351	VIRB type IV secr
24	33	56.9	319	2 T36845	probable membrane
25	33	56.9	335	2 E70655	hypothetical prote
26	33	56.9	353	2 C70502	hypothetical prote
27	33	56.9	366	2 C70365	twitching motility
28	33	56.9	425	2 TC4184	prolyl aminopeptid
29	33	56.9	445	2 T35893	FAD-dependent oxid

30	33	56.9	477	2 S68367	protoporphyrinogen
31	33	56.9	477	2 S65684	protoporphyrinogen
32	33	56.9	477	2 A56449	protoporphyrinogen
33	33	56.9	511	2 T40334	hypothetical prote
34	33	56.9	589	2 G87485	hypothetical prote
35	33	56.9	636	1 VCWPS	env polyprotein -
36	33	56.9	733	2 C97510	hypothetical prote
37	33	56.9	753	2 AB2729	succinoglycan bios
38	33	56.9	858	2 AE0999	penicillin-binding
39	33	56.9	1137	2 T19414	hypothetical prote
40	33	56.9	1940	2 F75393	hypothetical prote
41	32	55.2	57	2 I35055	MHC class II histo
42	32	55.2	136	2 H70543	hypothetical prote
43	32	55.2	149	2 AH2262	hypothetical prote
44	32	55.2	150	2 S09872	hypothetical prote
45	32	55.2	154	2 T34825	hypothetical prote

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M., Ozawa, A., Ishii, S., Shibusawa, N., Hashida, T., Ishizuka, T., Hosoya, B.  
Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior  
release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 93.1%; Score 54; DB 2; Length 83;  
Best Local Similarity 83.3%; Pred. No. 0.0015;  
Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRFX 12  
Db 41 YTGGRIRPVGRFX 52

RESULT 2  
E83663  
glutamyl-tRNA synthetase gltx [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C:Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
C:Accession: E83663  
R:Takami, H., Nakasone, K., Takaki, Y., Maeno, G., Sasaki, R., Masui, N., Fujii, F., H  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans a  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: E83663  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-485 <STO>  
A:Cross-references: GB:AP001507; GB:BA000004; MID:910172612; PIDN:BA03828.1; GSPDB:G  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: gltx  
C:Superfamily: glutamate--tRNA ligase; glutamine--tRNA ligase homology

Query Match 67.2%; Score 39; DB 2; Length 485;  
Best Local Similarity 58.3%; Pred. No. 8.8;  
Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| :||| :||  
DB 146 YKAKGIRKPVGRF 157

RESULT 3  
NEBP37  
endodeoxyribonuclease I (EC 3.1.21.-) - phage T7  
C:Species: phage T7  
C>Date: 01-Sep-1981 #sequence\_revision 24-Sep-1981 #text\_change 11-Jun-1999  
C:Accession: B94615; C92866; S42301; A00785  
R:Dunn, J.J.; Thompson, K.  
submitted to the Nucleic Acid Sequence Database, September 1982  
A:Reference number: A94615  
A:Accession: B94615  
A:Molecule type: DNA  
A:Residues: 1-149 <DU1>  
R:Dunn, J.J.; Studier, F.W.  
J. Mol. Biol. 148, 303-330, 1981  
A>Title: Nucleotide sequence from the genetic left end of bacteriophage T7 DNA to the be  
A:Reference number: A92866; MUID:82078034  
A:Accession: C92866  
A:Molecule type: DNA  
A:Residues: 1-149 <DU2>  
A:Cross-references: GB:V01127; NID:g15498; PIDN:CAA24345.1; PID:g15517  
J:Dunn, J.J.; Studier, F.W.  
J. Mol. Biol. 166, 477-535, 1983  
A>Title: Complete nucleotide sequence of bacteriophage T7 DNA and the locations of T7 ge  
A:Reference number: S42285; MUID:83241725  
A:Accession: S42301  
A:Molecule type: DNA  
A:Residues: 1-149 <DUN>  
A:Cross-references: EMBL:V01146; NID:g431187; PIDN:CAA24402.1; PID:g15581  
A>Note: the authors did not translate the codon for residue 1  
C:Comment: Endodeoxyribonuclease I, which is expressed in the late stage, is necessary f  
a linear monomer. In the late stage, the T7 DNA replicates via linear concatamers seve  
C:Genetics:  
A:Gene: 3  
A:Map position: 25.64-26.76  
C:Superfamily: phage T7 exodeoxyribonuclease I  
C:Keywords: hydrolase

Query Match 63.8%; Score 37; DB 1; Length 149;  
Best Local Similarity 58.3%; Pred. No. 6.4;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| :||| :||  
DB 4 YKAKGIRKPVGRF 15

RESULT 4  
AB7448  
conserved hypothetical protein CCL602 [imported] - Caulobacter crescentus  
C:Species: Caulobacter crescentus  
C>Date: 20-Apr-2001 #sequence\_revision 20-Apr-2001 #text\_change 10-May-2001  
C:Accession: AB7448  
R:Nleeman, W.C.; Feldblyum, T.V.; Paulsen, I.T.; Nelson, K.E.; Eisen, J.; Heidelberg, J.  
B.; Laub, M.T.; DeBoy, R.T.; Dodson, R.J.; Durkin, A.S.; Gwinn, M.L.; Haft, D.H.; Kolon  
N. J.; Esmolaeva, M.; White, O.; Salzberg, S.L.; Shapiro, L.; Venter, J.C.; Fraser, C.M.  
Proc. Natl. Acad. Sci. U.S.A. 98, 4136-4141, 2001  
A>Title: Complete Genome Sequence of Caulobacter crescentus.  
A:Reference number: AB7249; MUID:21173698; PMID:11259647  
A:Accession: AB7448  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-545 <STO>  
A:Cross-references: GB:AB005673; NID:g13422999; PIDN:AAK23581.1; GSPDB:GN00148

C:Genetics:  
A:Gene: CCL602  
C:Superfamily: hypothetical protein u1937b

Query Match 62.1%; Score 36; DB 2; Length 545;  
Best Local Similarity 50.0%; Pred. No. 39;  
Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| :||| :||  
DB 245 YRACGVPRVGR 256

RESULT 5  
G70875  
probable oxidoreductase (EC 1.5.99.-) - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: G70875  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
A.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:98295987  
A:Accession: G70875  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-674 <COL>  
A:Cross-references: GB:AL010186; GB:AL123456; NID:g3261493; PIDN:CAA15852.1; PID:g269  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: fadH  
C:Superfamily: Methylotrophus methylotrophus W3A1 trimethylamine dehydrogenase  
C:Keywords: 4Fe-4S; metalloprotein; oxidoreductase  
F:337,340,344,356/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 62.1%; Score 36; DB 2; Length 674;  
Best Local Similarity 85.7%; Pred. No. 49;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 GIRPVGR 11  
| :||| :||  
DB 492 GVRPVGR 498

RESULT 6  
T47959  
hypothetical protein F15G16.60 - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C>Date: 20-Apr-2000 #sequence\_revision 20-Apr-2000 #text\_change 20-Apr-2000  
C:Accession: T47959  
R:De Haan, M.; Maarse, A.C.; Grivell, L.A.; Nemes, H.W.; Lemcke, K.; Mayer, K.F.X.; O  
submitted to the Protein Sequence Database, January 2000  
A:Reference number: Z24480  
A:Accession: T47959  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-790 <DEH>  
A:Cross-references: EMBL:AL123959  
A:Experimental source: cultivar Columbia; BAC clone F15G16  
C:Genetics:  
A:Map position: 3  
A:introns: 39/1; 678/2; 698/3; 773/2  
A>Note: F15G16.60

Query Match 62.1%; Score 36; DB 2; Length 790;  
Best Local Similarity 50.0%; Pred. No. 57;  
Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;



OY 1 YXXRGIRPVGRF 12  
| : : : : |  
Db 373 YGSRGLQPHGRW 384

RESULT 7  
D95026  
ribosomal protein L6 [Imported] - Streptococcus pneumoniae (strain TIGR4)  
C:Species: Streptococcus pneumoniae  
C>Date: 03-Aug-2001 #sequence\_revision 03-Aug-2001 #text\_change 24-Aug-2001

C:Accession: D95026  
R:Retellin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heidon, J.D.; Mayhew, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtapple, nson, T.; Hickey, E.K.; Holt, I.E.  
Science 293, 498-506, 2001

A:Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a Virulent Isolate of Streptococcus pneumoniae.  
A:Reference number: A95000; MUID:21357209; PMID:11463916

A:Accession: D95026  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1178 <RUR>

A:CROSS-references: GB:AE005672; PIDN:AAK74405.1; PID:q14971694; GSPDB:GN00164; TIGR:SP4  
A:Experimental source: strain TIGR4  
C:Genetics:

A:Gene: SP0225  
C:Superfamily: Escherichia coli ribosomal protein L6

Query Match 60.3%; Score 35; DB 2; Length 178;  
Best Local Similarity 58.3%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| : : : : |  
Db 157 YGSRGLQPHGRW 168

RESULT 8  
D97897  
50S ribosomal protein L6 [Imported] - Streptococcus pneumoniae (strain R6)  
C:Species: Streptococcus pneumoniae  
C>Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 02-Nov-2001

C:Accession: D97897  
R:Roskams, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; Dehoff, B.S.; Ee, R.; Leblanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M  
y, P.; Sun, P.M.; Winkler, M.E.  
J. Bacteriol. 183, 5709-5717, 2001

A:Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;  
A:Title: Genome of the Bacterial Streptococcus pneumoniae Strain R6.  
A:Reference number: A97872; MUID:21429245; PMID:11544234

A:Accession: D97897  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1178 <RUR>

A:CROSS-references: GB:AE007317; PIDN:AAK99008.1; PID:q15457749; GSPDB:GN00174  
C:Genetics:

A:Gene: rplP  
C:Superfamily: Escherichia coli ribosomal protein L6

Query Match 60.3%; Score 35; DB 2; Length 178;  
Best Local Similarity 58.3%; Pred. No. 19;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
| : : : : |  
Db 157 YGSRGLQPHGRW 168

RESULT 9  
T39533

hypothetical protein SPBC16A3.19 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39533  
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
submitted to the EMBL Data Library, February 1998

A:Accession: T39533  
A:Reference number: 221862  
A:Status: preliminary; translated from GB/EMBL/DDAY

A:Molecule type: DNA  
A:Residues: 1-272 <MOO>  
A:CROSS-references: EMBL:AL021748; PIDN:CAA16870.1; GSPDB:GN00067; SPDB:SPBC16A3.19

A:Experimental source: strain 972h-; cosmid c16A3  
C:Genetics:

A:Gene: SPDB:SPBC16A3.19  
A:Map position: 2  
A:introns: 53/1; 97/3

Query Match 60.3%; Score 35; DB 2; Length 272;  
Best Local Similarity 85.7%; Pred. No. 30;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 RGRPVG 10  
| : : : : |  
Db 47 RGLRPVG 53

RESULT 10  
XXGZAC  
acetyl-CoA C-acetyltransferase (EC 2.3.1.9) - Zoogloea ramifera  
N:Alternate names: acetoacetyl-CoA thiolase; biosynthetic thiolase; thiolase II  
C:Species: Zoogloea ramifera  
C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 05-May-2000

C:Accession: A26121  
R:Peoples, O.P.; Masamune, S.; Walsh, C.T.; Sinskey, A.J.  
J. Biol. Chem. 262, 97-102, 1987

A:Title: Biosynthetic thiolase from Zoogloea ramifera. III. Isolation and characteriz  
A:Reference number: A26121; MUID:87083504  
A:Accession: A26121

A:Molecule type: DNA  
A:Residues: 1-391 <PEO>  
A:CROSS-references: EMBL:J02631; NID:q155617; PIDN:AAK27706.1; PID:q155618

A:Experimental source: strain I-16-M, ATCC 19623  
C:Comment: The active enzyme, a tetramer of identical chains, catalyzes the reversibl  
heats, and poly-beta-hydroxybutyrate biosynthesis.

C:Genetics:  
A:Gene: pnda  
C:Superfamily: acetyl-CoA acetyltransferase  
C:Keywords: acyltransferase; coenzyme A; homotetramer; ketone body metabolism; poly-b  
F;2-391/Product: acetyl-CoA acetyltransferase [status predicted <MAT>  
F;89/Active site: Cys #status predicted

Query Match 60.3%; Score 35; DB 1; Length 391;  
Best Local Similarity 75.0%; Pred. No. 43;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 RGRPVG 11  
| : : : : |  
Db 266 RGIQPLGR 273

RESULT 11  
B70821  
hypothetical protein RV0982 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70821

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garler, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998

A:Authors: Squires, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete genome  
A:Reference number: A70500; MUID:98295987  
A:Accession: B70821  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-504 <COI>  
A:Cross-references: GB:AL021999; GB:AL123456; NID:g3261538; PIDN:CA17581.1; PID:g291694  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV0982  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 60.3%; Score 35; DB 2; Length 504;  
Best Local Similarity 85.7%; Pred. No. 57;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 GIRPVR 11  
|:|||||  
DB 188 GIRPVR 194

RESULT 12  
745447  
Probable two-component system sensor [imported] - *Mycobacterium leprae*  
C:Species: *Mycobacterium leprae*  
C:Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
A:Accession: T45447  
R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
Submitted to the EMBL Data Library, February 1998  
A:Reference number: 222967  
A:Accession: T45447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-519 <JAM>  
A:Cross-references: EMBL:AL035500; PIDN:CAB36689.1  
A:Experimental source: cosmid L373  
C:Genetics:  
A:Note: MLCB373.27  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 60.3%; Score 35; DB 2; Length 519;  
Best Local Similarity 85.7%; Pred. No. 58;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 GIRPVR 11  
|:|||||  
DB 188 GIRPVR 194

RESULT 13  
E82221  
DNA gyrase, chain A VC1258 [imported] - *Vibrio cholerae* (strain N16961 serogroup O1)  
C:Species: *Vibrio cholerae*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
A:Accession: E82221  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Chadson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Drygol, I.; Sellers, J.  
L. R.; Melnikos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: AB2035; MUID:20406833  
A:Accession: E82221  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-694 <HEI>  
A:Cross-references: GB:AE004205; GB:AE003852; NID:g9655740; PIDN:AAF94417.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1258  
A:Map position: 1

C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomeras  
Query Match 60.3%; Score 35; DB 2; Length 894;  
Best Local Similarity 62.5%; Pred. No. 1e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 4 RGIRPVR 11  
|:|||||  
DB 733 KGRPVR 740

RESULT 14  
B8257  
Hypothetical protein XF2445 [imported] - *Xylella fastidiosa* (strain 9a5c)  
C:Species: *Xylella fastidiosa*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
A:Accession: B8257  
R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A:Reference number: AB2515; MUID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B8257  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1056 <SIM>  
A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF5244.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.  
Britones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carrero, D.M.; Carrer  
as-Neto, E.; Docena, C.; El-Deiry, H.; Facincini, A.P.; Ferreira, A.J.S.  
Submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Fr  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.  
Rodrigues, V.; Rosa, A.J. de M.; de Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeiri,  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tshako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovsky-Almeida, S.; Vettore, A.L.  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2445

Query Match 60.3%; Score 35; DB 2; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 GIRPVR 11  
|:|||||  
DB 1047 GIRPVR 1053

RESULT 15  
T36248  
CVA peptidase Streptomyces I - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 01-Dec-2000  
A:Accession: T36248  
R:Saunders, D.C.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream,  
submitted to the EMBL Data Library, March 1999  
A:Reference number: Z21602  
A:Accession: T36248  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-7463 <SAU>  
A:Cross-references: EMBL:AL035640; PIDN:CAB38518.1; GSPDB:GN00070; SCOEDB:SCE63.03c  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: cdaps1; SCOEDB:SCE63.03c

C:Superfamily: acetate--CoA ligase homology; acyl carrier protein homology  
 C:keywords: carrier protein; phosphopantetheine; phosphoprotein  
 F:516-1074/Domain: acetate--CoA ligase homology #status atypical <ACLI1>  
 F:1090-1158/Domain: acyl carrier protein homology <ACP1>  
 F:1715-2184/Domain: acetate--CoA ligase homology <ACLI2>  
 F:2200-2268/Domain: acyl carrier protein homology <ACP2>  
 F:2804-3249/Domain: acetate--CoA ligase homology <ACLI3>  
 F:3265-3332/Domain: acyl carrier protein homology <ACP3>  
 F:4323-4746/Domain: acetate--CoA ligase homology <ACLI4>  
 F:4762-4830/Domain: acyl carrier protein homology <ACP4>  
 F:5363-5786/Domain: acetate--CoA ligase homology <ACLI5>  
 F:5802-5870/Domain: acyl carrier protein homology <ACP5>  
 F:6401-6868/Domain: acetate--CoA ligase homology <ACLI6>  
 F:6884-6951/Domain: acyl carrier protein homology <ACP6>  
 F:1122,2232,3297,4794,5834,6916/Binding site: phosphopantetheine (Ser) (covalent) #statu

Query Match 60.38; Score 35; DB 2; Length 7463;  
 Best Local Similarity 63.68; Pred. No. 9.4e+02;  
 Matches 7; Conservative 0; Mismatches 4; Indels 0; Gaps 0;  
 OY 1 YXRGIRPVGR 11  
 | | | | |  
 Db 6680 YEARGSLPVGR 6690

Search completed: September 13, 2002, 09:24:02  
 Job time: 777 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:47 : Search time 80.21 Seconds  
(without alignments)  
6.275 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_9\_21  
Perfect score: 58  
Sequence: 1 YXXRCIRPVGRFX 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	55	94.8	87	1 PRRP_HUMAN	P81277 homo sapien
2	54	93.1	83	1 PRRP_RAT	P81278 ratulus norv
3	54	93.1	98	1 PRRP_BOVIN	P81264 bos taurus
4	39	67.2	485	1 SYE_BACHD	O9kgf6 bacillus ha
5	37	63.8	149	1 ENRN_BPT7	P00641 bacterioph
6	37	63.8	676	1 EXLI_HUMAN	O92935 homo sapien
7	35	60.3	179	1 RK6_GUTRH	O46908 gulliartha
8	35	60.3	391	1 THIL_ZOORA	P07037 zoogloea ra
9	35	60.3	922	1 GYRA_AERSA	P48359 aeromonas s
10	34	58.6	224	1 YGCI_ECOLI	O46898 escherichia
11	34	58.6	306	1 RSI_SYNP6	P46238 synechococ
12	34	58.6	328	1 RSLA_SYNY3	P73530 synechocyst
13	34	58.6	801	1 TFR2_HUMAN	O9up52 homo sapien
14	33	56.9	424	1 PIP_AERSO	P46347 aeromonas s
15	33	56.9	477	1 PPOX_HUMAN	P50336 homo sapien
16	33	56.9	477	1 PPOX_MOUSE	P51175 mus musculu
17	33	55.2	636	1 ENV_MCFE	P15073 mink cell f
18	32	55.2	150	1 ULAF_HCMVA	O55578 human cytom
19	32	55.2	249	1 Y361_SYNY3	O55578 synechocyst
20	32	55.2	330	1 AR72_HUMAN	O53468 homo sapien
21	32	55.2	331	1 AR73_HUMAN	O53154 homo sapien
22	32	55.2	433	1 ZU01_YEAST	P32527 saccharomyc
23	32	55.2	781	1 PBPB_HAEN	P45345 h penicillii
24	32	55.2	850	1 PBPB_ECOLI	P02918 escherichia
25	31	53.4	130	1 R13_PIG	O29293 sus scrofa
26	31	53.4	207	1 YDQ5_SCHPO	O14198 schizosacch
27	31	53.4	230	1 PYRE_HAEN	P43812 haemophilus
28	31	53.4	233	1 PYRE_PASU	O9cmml pseudoeilla
29	31	53.4	250	1 LINC_PSEPA	O9cmml pseudoeilla
30	31	53.4	364	1 MK11_MOUSE	O9wul1 mus musculu
31	31	53.4	367	1 GCST_MYCTU	O10376 mycobacteri
32	31	53.4	372	1 MK11_HUMAN	O15759 homo sapien
33	31	53.4	409	1 ENVI_FRSEV	P03393 itrend sple

34	31	53.4	413	1	CSD_THENA	O9x191 thermotoga
35	31	53.4	430	1	PYRC_MYCTU	P71809 mycobacteri
36	31	53.4	463	1	K10_DROME	P13468 drosophila
37	31	53.4	625	1	DHGL_DROME	P18173 drosophila
38	31	53.4	625	1	DHGL_DROPS	P18172 drosophila
39	31	53.4	631	1	DNAL_THERO	P66133 thermomicro
40	31	53.4	719	1	NRP1_YEAST	P32770 saccharomyc
41	31	53.4	733	1	SUF_DROME	P25991 drosophila
42	31	53.4	908	1	ACON_BACSU	P09339 bacillus su
43	31	53.4	988	1	SYI_SYNY3	P73505 synechocyst
44	31	53.4	1056	1	FOL_BIV06	P73560 bovine immu
45	31	53.4	1056	1	POL_BIV27	P19561 bovine immu

## ALIGNMENTS

```

RESULT 1
PRRP_HUMAN STANDARD: PRT: 87 AA.
AC P81277;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrRP31; Prolactin-
DE releasing peptide PrRP20].
GN PRP.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_Taxid=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98266781; Pubmed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsuno H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1996).
RN [2]
RX TISSUE SPECIFICITY.
RX Pubmed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
CC EMBL: AB015419; BAA29027.1; -
CC DR MIM: 602663; -
CC KW Hormone; Amidation; Signal.
CC FT SIGNAL 1 22
CC FT PEPTIDE 23 53 BY SIMILARITY.
CC FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP31.
CC FT MOD RES 53 53 PROLACTIN-RELEASING PEPTIDE PRRP20.
CC FT SEQUENCE 87 AA: 9639 MW: 229k23F50CF981B CRC64;
SQ

```

Query Match 94.8%; Score 55; DB 1; Length 87;  
 Best Local Similarity 83.3%; Pred. No. 0.00023;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
 DB 42 YASRGIRPVGRF 53

## RESULT 2

PRRP\_RAT STANDARD; PRT; 83 AA.  
 AC P81278;  
 DT 30-MAY-2000 (rel. 39, Created)  
 DT 30-MAY-2000 (rel. 39, Last sequence update)  
 DT 01-MAR-2002 (rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].

OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.,  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 RN [2]  
 RP TISSUE SPECIFICITY.

Pubmed=10498338;

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor.";  
 RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AB015418; BAA29026.1;  
 KW Hormone; Amidation; Signal;  
 FT SIGNAL 1 21  
 FT PEPTIDE 22 52  
 FT PEPTIDE 33 52  
 FT MOD\_RES 52 52  
 SQ SEQUENCE 83 AA; 9215 MW; DDC75A264EE4F29 CRC64;

Query Match 93.1%; Score 54; DB 1; Length 83;  
 Best Local Similarity 83.3%; Pred. No. 0.00035;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
 DB 41 YTGGRIRPVGRF 52

RESULT 3  
 PRRP\_BOVIN STANDARD; PRT; 98 AA.  
 ID P81264;  
 DT 30-MAY-2000 (rel. 39, Created)  
 DT 30-MAY-2000 (rel. 39, Last sequence update)  
 DT 01-MAR-2002 (rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovinae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.,  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AB015417; BAA29025.1;  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 FT SIGNAL 1 22  
 FT PEPTIDE 23 53  
 FT PEPTIDE 33 53  
 FT MOD\_RES 53 53  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA508 CRC64;

Query Match 93.1%; Score 54; DB 1; Length 98;  
 Best Local Similarity 83.3%; Pred. No. 0.00042;  
 Matches 10; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
 DB 42 YACRGIRPVGRF 53

## RESULT 4

SYE\_BACHD STANDARD; PRT; 485 AA.  
 ID SYE\_BACHD  
 AC O9KGF6;  
 DT 16-OCT-2001 (rel. 40, Created)  
 DT 16-OCT-2001 (rel. 40, Last sequence update)  
 DT 16-OCT-2001 (rel. 40, Last annotation update)  
 DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase) (GLURS).

DE GLURS;  
 GN GLTX OR BR0109.

OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 CC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=86665;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN-C-125 / JCM 9153;  
 RX MEDLINE-20512582; PubMed-11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kohara S.,  
 RA Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium *Bacillus*  
 RT *halodurans* and genomic sequence comparison with *Bacillus subtilis*.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) -> AMP +  
 CC diphosphate + L-glutamyl-tRNA(Glu).  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AP001507; BAB03828.1; -;  
 DR InterPro: IPR000924; tRNA-synt\_1c.  
 DR InterPro: IPR001412; tRNA-synt\_1.  
 DR PIR: PF00749; tRNA-synt\_1c; 1.  
 DR PRINTS: PR00987; TRNASYNTHGLU.  
 DR PROSITE: PS00178; AA.TRNA.LIGASE.1; 1.  
 KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KM Complete proteome.  
 FT SITE 11 21 \*HIGH\* REGION.  
 FT SITE 252 256 \*KMSKS\* REGION.  
 FT BINDING 255 255 ATP (BY SIMILARITY).  
 SO SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match 67.28; Score 39; DB 1; Length 485;  
 Best Local Similarity 58.3%; Pred. NO. 2.3;  
 Matches 7; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

OY 1 YXXGIRPVGRF 12  
 DB 146 YEAGKIRPVGRF 157

RESULT 5  
 ENRN\_BPT7 STANDARD: PRT: 149 AA.  
 AC P00641;  
 DT 21-JUL-1986 (Rel. 01, Created)  
 DT 21-JUL-1986 (Rel. 01, Last sequence update)  
 DT 01-NOV-1995 (Rel. 32, Last annotation update)  
 DE Endonuclease I (EC 3.1.21.2) (Endonuclease).  
 GN 3.  
 OS Bacteriophage T7.  
 OC Viruses; dsDNA viruses, no RNA stage; Caudovirales; Podoviridae;  
 OC T7-like phages.  
 OX NCBI\_TaxID-10760;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-83241725; PubMed-6864790;  
 RA Dunn J.J., Studier F.W.;  
 RT "Complete nucleotide sequence of bacteriophage T7 DNA and the  
 RT locations of T7 genetic elements.";  
 RL J. Mol. Biol. 166:477-535(1983).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-82078034; PubMed-7310871;  
 RA Dunn J.J., Studier F.W.;  
 RT "Nucleotide sequence from the genetic left end of bacteriophage T7  
 RT DNA to the beginning of gene 4.";  
 RL J. Mol. Biol. 148:303-330(1981).  
 CC -1- FUNCTION: ENDODEOXYRIBONUCLEASE I, WHICH IS EXPRESSED IN THE LATE

CC STAGE. IS NECESSARY FOR T7 GENETIC RECOMBINATION AND THE BREAKDOWN  
 CC OF HOST DNA. IN THE EARLY STAGE OF INFECTION, T7 DNA REPLICATES AS  
 CC A LINEAR MONOMER. IN THE LATE STAGE, THE T7 DNA REPLICATES VIA  
 CC LINEAR CONCATENERS SEVERAL GENOMES IN LENGTH. THE GENE 3 PRODUCT  
 CC HAS ALSO BEEN IMPLICATED IN THE MATURATION OF THESE CONCATENERS.  
 CC -1- CATALYTIC ACTIVITY: Endonucleolytic cleavage to 5'-  
 CC phosphooligonucleotide end-products.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: V01127; CA24345.1; -;  
 DR EMBL: V01146; CA24402.1; -;  
 DR PIR: A00785; NEBP37.  
 DR PIR: S42301; S42301.  
 KM Hydrolyase; Nuclease; Endonuclease.  
 SO SEQUENCE 149 AA; 17172 MW; D092AA28E3743BC1 CRC64;

Query Match 63.88; Score 37; DB 1; Length 149;  
 Best Local Similarity 58.3%; Pred. NO. 1.8;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXXGIRPVGRF 12  
 DB 4 YGAGKIRPVGRF 15

RESULT 6  
 EXTL\_HUMAN STANDARD: PRT: 676 AA.  
 AC Q92935;  
 DT 15-JUL-1998 (Rel. 36, Created)  
 DT 15-JUL-1998 (Rel. 36, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Exostosin-like 1 (Exostosin-L) (Multiple exostosin-like protein).  
 GN EXTL1 OR EXTL.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OX NCBI\_TaxID-9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97189339; PubMed-9037597;  
 RA Wise C.A., Clines G.A., Massa H., Trask B.J., Lovett M.;  
 RT "Identification and localization of the gene for EXTL, a third member  
 RT of the multiple exostosins gene family.";  
 RL Genome Res. 7:10-16(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Xu L., Deng H.X., Xia J.H., Pan Q., Liu C.Y.;  
 RT "Mutations of the EXT genes in hereditary multiple exostoses in  
 RT Chinese.";  
 RL Submitted (AUG-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA Wuyts M., Spleker N., Van Roy N., De Paep A., De Bouille K.,  
 RA Williams P.J., Van Hul W., Versteeg R., Speleman F.;  
 RT "Refined physical mapping and genomic structure of the EXTL1 gene.";  
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: APPEARS TO BE A TUMOR SUPPRESSOR (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. Endoplasmic  
 CC reticulum (BY similarity).  
 CC -1- SIMILARITY: BELONGS TO THE EXOSTOSIN FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: U67191; AAC51141.1; -  
 DR EMBL: AF083633; AAD02840.1; -  
 DR EMBL: AF083623; AAD02840.1; JOINED.  
 DR EMBL: AF083624; AAD02840.1; JOINED.  
 DR EMBL: AF083625; AAD02840.1; JOINED.  
 DR EMBL: AF083626; AAD02840.1; JOINED.  
 DR EMBL: AF083627; AAD02840.1; JOINED.  
 DR EMBL: AF083628; AAD02840.1; JOINED.  
 DR EMBL: AF083629; AAD02840.1; JOINED.  
 DR EMBL: AF083630; AAD02840.1; JOINED.  
 DR EMBL: AF083631; AAD02840.1; JOINED.  
 DR EMBL: AF083632; AAD02840.1; JOINED.  
 DR EMBL: AF153980; AAF73172.1; -  
 DR EMBL: AF151391; AAF73172.1; JOINED.  
 DR MIM: 601738; -  
 DR InterPro: IPR004263; Exostosin.  
 DR Pfam: PF03016; Exostosin\_1.  
 DR Anll-oncogene; Multigene family: Transmembrane; Signal-anchor.  
 FT TRANSMEM 10 30 SIGNAL-ANCHOR (Type-II MEMBRANE PROTEIN)  
 FT CAROXYD 269 269 (POTENTIAL)  
 FT SEQUENCE 676 AA; 74673 MW; B5E006A8762E5633 CRC64;  
 SQ  
 Query Match 63.8%; Score 37; DB 1; Length 676;  
 Best Local Similarity 58.3%; Pred. No. 8.4;  
 Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
 QY 1 YXXRGIRPYGRF 12  
 Db 408 YLQGSRGPRGRF 419  
 | : | | | | |  
 RESULT 7  
 RK6\_GUITH STANDARD: PRT; 179 AA.  
 AC 046908;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE Chloroplast 50S ribosomal protein L6.  
 GN RPL6.  
 OS Guillardia theta (Cryptomonas phl).  
 OC Chloroplast.  
 OC Eukaryota; Cryptophyta; Cryptomonadaceae; Guillardia.  
 OX NCBI\_TaxID=55529;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-9728375; PubMed-9137835;  
 RA Wang S.L., Liu X.-O., Douglas S.E.;  
 RT "The large ribosomal protein gene cluster of a cryptomonad plastid:  
 RT gene organization, sequence and evolutionary implications.";  
 RL Biochem. Mol. Biol. Int. 41:1035-1044(1997).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-9912821; PubMed-9929392;  
 RA Douglas S.E., Penny S.L.;  
 RT "The plastid genome of the cryptophyte alga, Guillardia theta:  
 RT complete sequence and conserved syntenic groups confirm its common  
 RT ancestry with red algae.";  
 RL J. Mol. Evol. 48:236-244(1999).  
 CC -1- FUNCTION: THIS PROTEIN BINDS DIRECTLY TO 23S RIBOSOMAL RNA AND IS  
 CC LOCATED AT THE AMINOACYL-TRNA BINDING SITE OF THE  
 CC PEPTIDYLTRANSFERASE CENTER (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO THE L6P FAMILY OF RIBOSOMAL PROTEINS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AF041468; AAC35717.1; -  
 DR HSSP: P02391; 1RL6.  
 DR InterPro: IPR000702; Ribosomal\_L6.  
 DR InterPro: IPR002358; Ribosomal\_L6\_1.  
 DR Pfam: PF00347; Ribosomal\_L6\_1.  
 DR PRINTS: PR00059; RIBOSOMAL\_L6.  
 DR PRODOM: PD002236; Ribosomal\_L6\_1.  
 DR PROSITE: PS00525; RIBOSOMAL\_L6\_1; 1.  
 KW Ribosomal protein; Chloroplast.  
 SQ SEQUENCE 179 AA; 19527 MW; 8B4C0BDB0152AD24 CRC64;  
 QY 1 YXXRGIRPYGRF 12  
 Db 157 YKKGIRVGEF 168  
 | : | | | | |  
 RESULT 8  
 THIL\_ZOORA STANDARD: PRT; 391 AA.  
 AC P07097;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Acetyl-CoA acetyltransferase (EC 2.3.1.9) (Acetoacetyl-CoA thiolase).  
 GN PHBA.  
 OS Zoogloea ramifera.  
 OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;  
 OC Zoogloea.  
 OX NCBI\_TaxID=350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-ATCC 19623 / I-16-M;  
 RX MEDLINE-87083504; PubMed-2878929;  
 RA Peoples O.P., Masamune S., Walsh C.T., Sinskey A.J.;  
 RT "Biosynthetic thiolase from Zoogloea ramifera. III. Isolation and  
 RT characterization of the genes encoding beta-ketothiolase and  
 RT acetoacetyl-CoA reductase.";  
 RL J. Biol. Chem. 264:15293-15297(1989).  
 RN [2]  
 RP REVISION NO 130.  
 RX MEDLINE-89359356; PubMed-2670935;  
 RA Peoples O.P., Sinskey A.J.;  
 RT "Poly-beta-hydroxybutyrate biosynthesis in Alcaligenes eutrophus H16.  
 RT characterization of the genes encoding beta-ketothiolase and  
 RT acetoacetyl-CoA reductase.";  
 RL J. Biol. Chem. 264:15293-15297(1989).  
 RN [3]  
 RP MUTAGENESIS OF CYS-377.  
 RX MEDLINE-91217075; PubMed-1673680;  
 RA Palmer M.A.J., Differding E., Gamboni R., Williams S.F., Peoples O.P.,  
 RA Walsh C.T., Sinskey S.J., Masamune S.;  
 RT "Biosynthetic thiolase from Zoogloea ramifera. Evidence for a  
 RT mechanism involving Cys-378 as the active site base.";  
 RL J. Biol. Chem. 266:8369-8375(1991).  
 CC -1- CATALYTIC ACTIVITY: 2 acetyl-CoA + acetoacetyl-CoA.  
 CC -1- PATHWAY: FIRST STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its



CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: J02631; AAA27706.1; ALT\_SEQ.  
DR PIR: A26121; XKCZAC.  
DR PIR: A27754; A27754.  
DR HSSP: P27796; IAFY.  
DR InterPro: IPR002155; Thiolase.  
DR Pfam: PF00108; thiolase\_1.  
DR Pfam: PF02803; thiolase\_C\_1.  
DR PROSITE: PS00098; THIOLEASE\_1; 1.  
DR PROSITE: PS00099; THIOLEASE\_2; 1.  
DR PROSITE: PS00737; THIOLEASE\_3; 1.  
DR Transferrase; Acyltransferase; PHB biosynthesis.  
KW TRANSFERASE; ACYLTRANSFERASE; PHB BIOSYNTHESIS.  
FT INIT\_MET 0 0 SUBSTRATE BINDING (BY SIMILARITY).  
FT ACT\_SITE 88 88 BASE.  
FT ACT\_SITE 377 377 C->G; LOSS OF ACTIVITY.  
FT MUTAGEN 377 377 C->G; LOSS OF ACTIVITY.  
SQ SEQUENCE 391 AA; 40342 MM; 6D2351A1BC0F4EDD CRC64;  
  
Query Match 60.3%; Score 35; DB 1; Length 391;  
Best Local Similarity 75.0%; Pred. No. 12;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
  
OY 4 RGIRPVGR 11  
|||:|:|  
Db 266 RGIRPVGR 273  
  
RESULT 9  
GYRA\_AERSA STANDARD; PRT; 922 AA.  
ID GYRA\_AERSA  
AC PA8369;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE DNA gyrase subunit A (EC 5.99.1.3).  
GN GYRA.  
OS Aeromonas salmonicida.  
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
CC Aeromonas.  
OX NCBI\_TaxID=645;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=2148/89;  
RA Opegaard H.;  
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 33-179 FROM N.A.  
RC STRAIN=ATCC 14174;  
RX MEDLINE=95142596; PubMed=7840589;  
RA Opegaard H.; Sorum H.;  
RT "gyrA mutations in quinolone-resistant isolates of the fish pathogen  
RT Aeromonas salmonicida.";  
RL Antimicrob. Agents Chemother. 38:2460-2464(1994).  
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE  
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED  
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
CC of double-stranded DNA.  
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA  
CC BREAKAGE AND REJOINING. THE B CHAIN CATALYZES ATP HYDROLYSIS. THE  
CC ENZYME FORMS AN A222 TETRAMER.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial

CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: L47978; AAB41037.1; -;  
DR EMBL: L42453; AAB87239.1; -;  
DR HSSP: P09097; IAB4.  
DR InterPro: IPR002205; DNA\_topoisomV.  
DR Pfam: PF00521; DNA\_topoisomV; 1.  
DR SMART: SM00434; TOPAC; 1.  
KW Isomerase; Topoisomerase; DNA-binding.  
FT ACT\_SITE 122 122 DNA CLEAVAGE (BY SIMILARITY).  
SQ SEQUENCE 922 AA; 101333 MM; 8894965DC4217077 CRC64;  
  
Query Match 60.3%; Score 35; DB 1; Length 922;  
Best Local Similarity 62.5%; Pred. No. 29;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
  
OY 4 RGIRPVGR 11  
|||:|:|  
Db 758 RGIRPVGR 765  
  
RESULT 10  
YGC1\_ECOLI STANDARD; PRT; 224 AA.  
ID YGC1\_ECOLI  
AC Q46898;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein ygc1 precursor.  
GN YGC1 OR B2757.  
OS Escherichia coli.  
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
CC Escherichia.  
OX NCBI\_TaxID=562;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=K12 / MG1655;  
RX MEDLINE=97426617; PubMed=9278503;  
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12.";  
RL Science 277:1453-1474(1997).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U29579; AAA69267.1; ALT\_INIT.  
DR EMBL: AE000359; AAC75799.1; ALT\_INIT.  
DR Ecogene: EG13116; YGC1.  
KW Hypothetical protein; Signal; Complete proteome.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 224 HYPOTHETICAL PROTEIN YGC1.  
SQ SEQUENCE 224 AA; 25209 MM; 1C42CC009B317D68 CRC64;  
  
Query Match 58.6%; Score 34; DB 1; Length 224;  
Best Local Similarity 75.0%; Pred. No. 11;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 5 GRRPVGRF 12  
||| |||  
Db 23 GRRPVGRF 30

```

RESULT 11
RS1_SYNP6 STANDARD; PRT; 306 AA.
AC P46228;
DT 01-NOV-1995 (Rel. 32, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S1.
GN RPSA OR RPSL.
OS Synechococcus sp. (strain PCC 6301) (Anacystis nidulans).
OC Bacteria; Cyanobacteria; Chroococcales; Synechococcus.
OX NCBI_TaxID=1139;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-13.
RX MEDLINE=9516170; PubMed=7862084;
RA Sugita M., Sugita C., Sugita M.;
RT "Structure and expression of the gene encoding ribosomal protein S1
RT from the cyanobacterium Synechococcus sp. strain PCC 6301: striking
RT sequence similarity to the chloroplast ribosomal protein S1.";
RL Mol. Gen. Genet. 246:142-147(1995).
CC -1- FUNCTION: BINDS MRNA.
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: CONTAINS 3 SI MOTIF DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D28752; BAA05946.1; -.
DR HSSP: P05055; ISRO.
DR InterPro: IPR003029; S1.
DR Pfam: PF00575; S1; 3.
DR SMART: SM00316; S1; 3.
DR PROSITE: PS50126; S1; 3.
KW Ribosomal protein; Repeat; RNA-binding.
FT INIT_MET 0
FT DOMAIN 31 100 SI MOTIF 1.
FT DOMAIN 118 182 SI MOTIF 2.
FT DOMAIN 196 264 SI MOTIF 3.
FT DOMAIN 306 34492 MW; 82DA4A2EBE6B97CE CRC64;
SO SEQUENCE

Query Match 58.6%; Score 34; DB 1; Length 306;
Best Local Similarity 66.7%; Pred. No. 15;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 RGIRPVGRF 12
DB 204 RGIRPVGRF 212

RESULT 12
RS1A_SYNP3 STANDARD; PRT; 328 AA.
AC P73530;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE 30S ribosomal protein S1 homolog A.
GN RPS1A OR SLR1356.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hirosewa M., Sugita M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K.,

```

```

RA Okumura S., Shimpou S., Takeuchi C., Wada T., Watanabe A.,
RA Yamada M., Yasuda M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136(1996).
CC -1- FUNCTION: BINDS MRNA.
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.
CC -1- SIMILARITY: CONTAINS 3 SI MOTIF DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL: D90907; BAA17570.1; -.
DR HSSP: P05055; ISRO.
DR InterPro: IPR003029; S1.
DR Pfam: PF00575; S1; 3.
DR SMART: SM00316; S1; 3.
DR PROSITE: PS50126; S1; 3.
KW Ribosomal protein; Repeat; RNA-binding; Complete proteome.
FT DOMAIN 31 100 SI MOTIF 1.
FT DOMAIN 118 182 SI MOTIF 2.
FT DOMAIN 196 264 SI MOTIF 3.
FT DOMAIN 328 36570 MW; DC3FF9B1E5A40619 CRC64;
SO SEQUENCE

Query Match 58.6%; Score 34; DB 1; Length 328;
Best Local Similarity 66.7%; Pred. No. 16;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 4 RGIRPVGRF 12
DB 204 RGIRPVGRF 212

RESULT 13
TFR2_HUMAN STANDARD; PRT; 801 AA.
AC Q9UB52; Q9H499; Q9N467;
DT 01-MAR-2002 (Rel. 41, Created)
DT 01-MAR-2002 (Rel. 41, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Transferrin receptor protein 2 (TFR2).
GN TFR2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A. (ISOFORM ALPHA AND BETA).
RC TISSUE=Erythrocytes, and Myeloid leukemia cells;
RX MEDLINE=9340005; PubMed=10409623;
RA Kawabata H., Yang R., Hirama T., Vuong P.T., Kawano S., Gombart A.F.,
RA Koefler H.P.;
RT "Molecular cloning of transferrin receptor 2: a new member of the
RT transferrin receptor-like family.";
RL J. Biol. Chem. 274:20826-20832(1999).
RN [2]
RP SEQUENCE FROM N.A. (ISOFORM GAMMA).
RX PubMed=9799793;
RA Gloeckner G., Scherer S., Schatteroy R., Boright A.P., Weber J.,
RA Tsui L.-C., Rosenthal A.;
RT "Large scale analysis of two regions in human chromosome 7q22:
RT annotation of 650 kb of genomic sequence around the EPO and CUTL1 loci
RT reveals 17 genes.";
RL Genome Res. 8:1060-1073(1998).
RN [3]
RP SEQUENCE OF 1-158 AND 370-801 FROM N.A.

```

RC TISSUE-Carcinoma, and Embryo; Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Isegawa T., Ota T., Hayashi K., Sugano S., Shiratori A., Sudo H.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Megatsuma M., Hosokawa T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
 RA Matsunabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuko Y.,  
 RA Niinomiya K., Iwayanagi T.;  
 RT "MEDO human cDNA sequencing project.";  
 RN Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
 RP DISBASE.  
 RX PubMed:10802645;  
 RA Camaschella C., Roetto A., Cali A., De Gobbi M., Garozzo G.,  
 RA Camaschella M., Majorano N., Totaro A., Gasparini P.;  
 RT "The gene TFR2 is mutated in a new type of haemochromatosis mapping to  
 7q22";  
 RL Nat. Genet. 25:14-15(2000).  
 RN [5]  
 RP VARIANT HFE3 Lys-172.  
 RX PubMed:11313241;  
 RA Roetto A., Totaro A., Piperno A., Piga A., Longo F., Garozzo G.,  
 RA Cali A., De Gobbi M., Gasparini P., Camaschella C.;  
 RT "New mutations inactivating transferrin 2 in hemochromatosis type  
 3";  
 RL Blood 97:2555-2560(2001).  
 CC -1- FUNCTION: Mediates cellular uptake of transferrin-bound iron in a  
 non-iron dependent manner. May be involved in iron metabolism,  
 hepatocyte function and erythrocyte differentiation.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. The beta isoform  
 lacks the transmembrane domain and is probably intracellular.  
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms: alpha (shown here), beta and  
 gamma: are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in liver. While the  
 alpha form is also expressed in spleen, lung, muscle, prostate and  
 peripheral blood mononuclear cells, the beta form is expressed in  
 all tissues tested, albeit weakly.  
 CC -1- DISEASE: Defects in TFR2 are a cause of hereditary hemochromatosis  
 type III (HFE3). HFE3 is a disorder of iron homeostasis resulting  
 in iron overload and has a phenotype indistinguishable from that  
 of hereditary hemochromatosis (HH). HH is characterized by  
 abnormal intestinal iron absorption and progressive increase of  
 total body iron, which results in midlife in clinical  
 complications including cirrhosis, cardiopathy, diabetes,  
 endocrine dysfunctions, arthropathy, and susceptibility to liver  
 cancer. Since the disease complications can be effectively  
 prevented by regular phlebotomies, early diagnosis is most  
 important to provide a normal life expectancy to the affected  
 subjects.  
 CC -1- MISCELLANEOUS: The variant Lys-172 found in hereditary  
 hemochromatosis type III affects the putative initiation codon of  
 the beta isoform thus preventing its translation.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: AF067864; AAD45561.1; -;  
 DR EMBL: AF053356; AAC78796.1; -;  
 DR EMBL: AK022002; BAB13951.1; -;  
 DR EMBL: AK000421; BAA91153.1; ALT\_INIT.  
 DR MEROPS: M28.973; -;  
 DR MIM: 604720; -;  
 DR MIM: 604250; -;  
 DR InterPro: IPR003137; PA.  
 DR Pfam: PF02225; PA; 1.  
 KW Transmembrane; Glycoprotein; Receptor; Signal-anchor;

KW Alternative splicing; Disease mutation.  
 FT DOMAIN 1 83 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 84 104 SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)  
 FT (POTENTIAL).  
 FT DOMAIN 105 801 EXTRACELLULAR (POTENTIAL).  
 FT SITE 23 26 ENDOCYTOSIS SIGNAL (POTENTIAL).  
 FT DISULFID 108 108 INTERCHAIN (POTENTIAL).  
 FT DISULFID 111 111 INTERCHAIN (POTENTIAL).  
 FT CARBOHYD 240 240 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 339 339 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 540 540 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT CARBOHYD 754 754 N-LINKED (GLCNAC. . .) (POTENTIAL).  
 FT VARSPLIC 343 369 MISSING (IN ISOFORM BETA).  
 FT VARSPLIC 172 172 MISSING (IN ISOFORM GAMMA).  
 FT VARIANT 172 172 M -> K (IN HFE3).  
 FT CONFLICT 712 712 /FtId-VAR.012738.  
 FT SEQUENCE 801 AA; 88755 MW; D3D3082BA835413A CRC64;  
 SO  
 Query Match 58.6%; Score 34; DB 1; Length 801;  
 Best Local Similarity 75.0%; Pred. No. 40;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 4 RGIRPVGR 11  
 Db 262 RGVDPVGR 269  
 ID PIP\_AERSO STANDARD; PRT; 424 AA.  
 AC P46547;  
 DT 01-NOV-1995 (Rel. 32, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Proline iminopeptidase (EC 3.4.11.5) (PIP) (Prolyl aminopeptidase)  
 DE (PAP).  
 GN PIP OR PAP.  
 OS Aeromonas sobria.  
 CC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
 CC Aeromonas.  
 CC NCBI\_TaxID=646;  
 CC [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=95189781; PubMed=7883756;  
 RA Kitazono A., Kitano A., Tsuru D., Yoshimoto T.;  
 RT "Isolation and characterization of the prolyl aminopeptidase gene  
 (pap) from Aeromonas sobria: comparison with the Bacillus coagulans  
 enzyme.";  
 RL J. Biochem. 116:818-825(1994).  
 CC -1- FUNCTION: HIGHER ACTIVITY TOWARD LONG PEPTIDES. ACTS ON  
 HYDROXYPROLINE BETA-NAPHTHYLAMIDE WITH ALMOST AS HIGH AN ACTIVITY  
 AS ON PROLINE BETA-NAPHTHYLAMIDE.  
 CC -1- CATALYTIC ACTIVITY: Release of a N-terminal proline from a  
 peptide.  
 CC -1- SUBUNIT: HOMOTETRAMER.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY S33.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 the European Bioinformatics Institute. There are no restrictions on its  
 use by non-profit institutions as long as its content is in no way  
 modified and this statement is not removed. Usage by and for commercial  
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL: D30714; BAA06380.1; -;  
 DR MEROPS: S33.001; -;  
 DR InterPro: IPR000073; Abhydrolase.  
 DR InterPro: IPR000379; Est\_11p\_Thioest\_actsite.  
 DR InterPro: IPR002410; Pro\_aminopeptidase.

DR pfam: PF00561; abhydrolase; 1.  
 DR PRINTS: PRO0793; PROAMNOPTASE.  
 KW Hydrolase; Amino-peptidase.  
 FT INIT\_MER 0 0  
 FT ACT\_SITE 145 145 NUCLEOPHILE (BY SIMILARITY).  
 FT ACT\_SITE 366 366 BY SIMILARITY.  
 FT ACT\_SITE 403 403 PROTON DONOR (BY SIMILARITY).  
 SO SEQUENCE 424 AA; 48274 MW; E9DD968E490047DC CRC64;

Query Match  
 Best Local Similarity 45.5%; Score 33; DB 1; Length 424;  
 Matches 5; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXRGIRPVGR 11  
 DB 166 YLTGVAPIGR 176

RESULT 15  
 ID PPOX\_HUMAN STANDARD; PRT; 477 AA.  
 AC P50336;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Protoporphyrinogen oxidase (EC 1.3.3.4) (PPO).  
 GN PPOX.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE=95229621; PubMed=713909;  
 RA Nishimura K., Taketani S., Inokuchi H.;  
 RT "Cloning of a human cDNA for protoporphyrinogen oxidase by  
 RT complementation in vivo of a hemg mutant of Escherichia coli.";  
 RL J. Biol. Chem. 270:8076-8080(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE=96367087; PubMed=8771201;  
 RA Dailey T.A., Dailey H.A.;  
 RT "Human protoporphyrinogen oxidase: expression, purification, and  
 RT characterization of the cloned enzyme.";  
 RL Protein Sci. 5:98-105(1996).  
 RN [3]  
 RP VARIANT VP ARG-232, AND VARIANT HIS-304.  
 RX MEDLINE=97005368; PubMed=8852667;  
 RA Deydach J.-C., Puy H., Robreau A.-M., Lamoril J., da Silva V.,  
 RA Grandchamp B., Nordmann Y.;  
 RT "Mutations in the protoporphyrinogen oxidase gene in patients with  
 RT variegate porphyria.";  
 RL Hum. Mol. Genet. 5:407-410(1996).  
 RN [4]  
 RP VARIANT VP TRP-59, AND VARIANT CYS-168.  
 RX MEDLINE=96241580; PubMed=8673113;  
 RA Melsner P.N., Dailey T.A., Hift R.J., Ziman M., Corrigan A.V.,  
 RA Roberts A.G., Melsner D.M., Kirsch R.E., Dailey H.A.;  
 RT "A R59W mutation in human protoporphyrinogen oxidase results in  
 RT decreased enzyme activity and is prevalent in South Africans with  
 RT variegate porphyria.";  
 RL Nat. Genet. 13:95-97(1996).  
 RN [5]  
 RP VARIANT VP CYS-152.  
 RX MEDLINE=98434271; PubMed=9763307;  
 RA Frank J., Poh-Fitzpatrick M.B., King L.E. Jr., Cristiano A.M.;  
 RT "The genetic basis of 'Scarsdale Gourmet Diet' variegate porphyria: a  
 RT missense mutation in the protoporphyrinogen oxidase gene.";  
 RL Arch. Dermatol. Res. 290:441-445(1998).  
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN

CC IX TO FORM PROTOPORPHYRIN IX.  
 CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) -> protoporphyrin-IX + H(2)O(2).  
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.  
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG, LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.  
 CC -1- DISEASE: DEFECTS IN PPOX ARE THE CAUSE OF PORPHYRIA VARIEGATA (VP), A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR HYPERTRICHOSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE INTERMITTENT PORPHYRIA.  
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC EMBL: D38537; BAA07538.1; -;  
 CC EMBL: U26446; AAA67690.1; -;  
 CC MIM: 600923; -;  
 CC DR MIM: 176200; -;  
 CC KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;  
 CC Flavo-protein; FAD; Mitochondrion; Disease mutation; Polymorphism.  
 CC FT NP\_BIND 9 14 FAD (POTENTIAL).  
 CC FT VARIANT 59 14 R -> W (IN VP).  
 CC FT VARIANT 152 152 /FTID-VAR\_003686.  
 CC FT VARIANT 168 168 R -> C (IN VP).  
 CC FT VARIANT 232 232 /FTID-VAR\_003687.  
 CC FT VARIANT 304 304 R -> H.  
 CC SO SEQUENCE 477 AA; 50765 MW; 2444DEAC26C33EE CRC64;

Query Match  
 Best Local Similarity 56.9%; Score 33; DB 1; Length 477;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 4 RGIRPVG 10  
 DB 59 RGIRPAG 65

Search completed: September 13, 2002, 09:30:47  
 Job time: 1137 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:22 ; Search time 311.85 Seconds  
(without alignments)  
7.212 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_9\_21  
Perfect score: 58  
Sequence: 1 YXXRGIRPVGRF 13

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteriophage:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	51	87.9	117	13 Q9W624	Q9W624 carassius a
2	43	74.1	420	2 Q93LZ7	Q93LZ7 streptomyce
3	40	69.0	153	9 Q9T133	Q9T133 bacterioph
4	37	63.8	348	5 Q9W1G4	Q9W1G4 drosophila
5	37	63.8	885	16 Q9ZNX4	Q9ZNX4 rhizobium m
6	37	63.8	1450	11 Q54728	Q54728 ratius norv
7	36	62.1	426	17 Q979S0	Q979S0 thermoplasm
8	36	62.1	545	16 Q9A7W7	Q9A7W7 caulobacter
9	36	62.1	674	16 Q50431	Q50431 mycobacteri
10	36	62.1	790	10 Q9M371	Q9M371 arabidopsis
11	35	60.3	144	4 Q96MP2	Q96MP2 homo sapien
12	35	60.3	178	16 Q97S07	Q97S07 streptococ
13	35	60.3	272	3 Q42924	Q42924 schizosacch
14	35	60.3	371	16 Q981V3	Q981V3 rhizobium 1
15	35	60.3	504	16 Q53895	Q53895 mycobacteri
16	35	60.3	519	16 Q9Z5G7	Q9Z5G7 mycobacteri

17	35	60.3	874	5 Q9VSE3	Q9VSE3 drosophila
18	35	60.3	894	16 Q9K5D8	Q9K5D8 vibrio chol
19	35	60.3	1000	5 Q9VS40	Q9VS40 drosophila
20	35	60.3	1056	16 Q9PA00	Q9PA00 xyella fas
21	35	60.3	1172	5 Q9VU79	Q9VU79 drosophila
22	35	60.3	7463	2 Q9Z4X6	Q9Z4X6 streptomyce
23	34	58.6	185	2 Q9XC09	Q9XC09 mycobacteri
24	34	58.6	278	16 Q9Z0R1	Q9Z0R1 rhizobium m
25	34	58.6	329	2 Q9RJR7	Q9RJR7 streptomyce
26	34	58.6	529	3 Q9C1M8	Q9C1M8 pleurotus s
27	34	58.6	537	4 Q75230	Q75230 homo sapien
28	34	58.6	780	4 Q75422	Q75422 homo sapien
29	34	58.6	801	4 Q9UP52	Q9UP52 homo sapien
30	33	56.9	68	16 Q9WKN3	Q9WKN3 thermotoga
31	33	56.9	139	5 Q9N6S1	Q9N6S1 leishmania
32	33	56.9	205	4 Q96TC9	Q96TC9 homo sapien
33	33	56.9	253	2 P72459	P72459 streptomyce
34	33	56.9	303	4 Q9BV19	Q9BV19 homo sapien
35	33	56.9	313	16 Q9ZY28	Q9ZY28 rhizobium m
36	33	56.9	319	2 Q88067	Q88067 streptomyce
37	33	56.9	335	16 P96221	P96221 mycobacteri
38	33	56.9	353	16 Q33194	Q33194 mycobacteri
39	33	56.9	366	16 Q66950	Q66950 aquifex aeo
40	33	56.9	440	15 Q85561	Q85561 mink cell f
41	33	56.9	445	2 Q50531	Q50531 streptomyce
42	33	56.9	477	11 Q99M34	Q99M34 mus musculu
43	33	56.9	511	3 Q59714	Q59714 schizosacch
44	33	56.9	589	16 Q9A7I8	Q9A7I8 caulobacter
45	33	56.9	636	15 Q85506	Q85506 murine leuk

## ALIGNMENTS

RESULT 1					
Q9W624	1	PRELIMINARY:	PRT:	117 AA.	
AC Q9W624:					
DT 01-NOV-1999 (TREMBLrel. 12, Created)					
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)					
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					
DE C-RR AMIDE.					
OS Carassius auratus (Goldfish).					
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;					
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;					
OC Cypriniformes; Cyprinidae; Carassius.					
OX NCBI_TaxID=7957;					
RN [1]					
RP SEQUENCE FROM N.A.					
RC TISSUE-BRAIN;					
RA Satake H., Minakata H., Fujimoto M.;					
RT "Carassius Rfam1de (C-RR amide)".					
RI Submitted (NOV-1998) to the EMBL/Genbank/DBJ databases.					
DR EMBL: AB020024; BAA76662.1; ..					
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;					
Query Match	87.9%	Score 51:	DB 13:	Length 117:	
Best Local Similarity	66.7%	Pred. No. 0.015;			
Matches 8;	Conservative 2;	Mismatches 2;	Indels 0;	Gaps 0;	
QY 1 YXXRGIRPVGRF 12					
DB 64 YVGRGVPRIGRF 75					
RESULT 2					
Q93LZ7	1	PRELIMINARY:	PRT:	420 AA.	
AC Q93LZ7:					
DT 01-DEC-2001 (TREMBLrel. 19, Created)					
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)					
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)					

OC Euteuryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephyraoidea; Drosophilidae; Drosophila.  
OX NCBI\_taxid=7227;  
[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-BERKELEY;  
RC MEDLINE=20196006; PubMed=10731132;  
RX  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Goebye J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brannon R.C., Rogers Y.-H.C., Blaise R.G., Champagne M., Pfeiffer B.D.,  
RA Man K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Mklos G.L.G.,  
RA Abell J.F., Aghayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
RA Beeson K.T., Benos P.V., Bertram B.P., Bhandari D., Bolshakov S.,  
RA Borikova D., Botchan M.R., Bouck J., Brokstein P., Broctier P.,  
RA Butrlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Daveport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Dudbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Garibaylan A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostali D., Houston K.A., Howland T.J., Wei M.-H., Ibeagwa C.,  
RA Jostali M., Koushun F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Klupp D., Lal Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mactel B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon R., Nuskern D.R., Paciel J.M.,  
RA Palazzolo M., Peltman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Sildenafil M., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*."  
RL Science 287:2185-2195(2000).  
RT -1 - SIMILARITY: CONSTRAINTS 7 WD REPEATS (TRP-ASP DOMAINS).  
CC EMBL: AE003462; AAF47100.1; -  
DR FlyBase: FBgn0034931; CG2812.  
DR InterPro: IPR001680; WD40.  
DR Pfam: PF00400; WD40\_7.  
DR PRINTS: PRO0320; GPROTEINBRPT.  
DR SMART: SM00320; WD40\_7.  
DR PROSITE: PS50082; WD\_REPEATS\_2; 4.  
DR PROSITE: PS50294; WD\_REPEATS\_REGION; 1.  
KW Repeat; WD repeat.  
SQ SEQUENCE 348 AA; 38787 MW; B001050A0E6F65A0 CRC64;

Query Match 63.8%; Score 37; DB 5; Length 348;  
Best Local Similarity 58.3%; Pred. No. 31;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;

RESULT 5  
ID 092NX4 PRELIMINARY; PRT. 885 AA.  
AC 092NX4;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)

DT 01-DEC-2001 (TREMblrel. 19, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE PTATIVE IRON/HEME TRANSPORT PROTEIN.  
 GN SMC04205.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 OC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=1021;  
 RX MEDLINE=2136823; PubMed=11474104;  
 RA Galibert F., Finan T.M., Long S.R., Puehler A., Abola P., Ampe F.,  
 RA Barloy-Hubler F., Barnett M.J., Becker A., Boistard P., Bothe G.,  
 RA Bouly M., Bowser L., Buhmester J., Cadieu E., Capela D., Chain P.,  
 RA Cowle A., Davis R.W., Dreano S., Federspiel N.A., Fisher R.F.,  
 RA Gloux S., Godrie T., Goffeau A., Golding B., Gouzy J., Gurjal M.,  
 RA Hernandez-Lucas I., Hong A., Huizar L., Hyman R.W., Jones T., Kahn D.,  
 RA Kahn M.L., Kaiman S., Keating D.H., Kiss E., Komp C., Lelaure V.,  
 RA Masuy D., Palm C., Peck M.C., Pohl T.M., Portetelle D., Purnelle B.,  
 RA Ramsperger U., Surzycki R., Thebaud P., Vandenberg M.,  
 RA Vorholter F.J., Weidner S., Wells D.H., Wong K., Yeh K.-C., Batut J.;  
 RT "The composite genome of the legume symbiont Sinorhizobium meliloti.";  
 RL Science 293:668-672(2001).  
 DR EMBL; AL591789; CAC46612.1; -.  
 KW Complete proteome.  
 SQ SEQUENCE 885 AA; 95389 MW; BEE3664106DDBF4 CRC64;

Query Match 63.8%; Score 37; DB 16; Length 885;  
 Best Local Similarity 54.5%; Pred. No. 86;  
 Matches 6; Conservative 2; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 YXXRGIRPVGR 11  
 Db 511 YGYKGLNPVGR 521

RESULT 6  
 ID 054728 PRELIMINARY; PRT; 1450 AA.  
 AC 054728;  
 DT 01-JUN-1998 (TREMblrel. 06, Created)  
 DT 01-JUN-1998 (TREMblrel. 06, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE PHOSPHOLIPASE B.  
 OS Rattus norvegicus (Rat).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 OX NCBI\_TaxID=10116;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SD; TISSUE=SMALL INTESTINE;  
 RX MEDLINE=98113187; PubMed=9442065;  
 RA Takemori H., Zolotariov F.N., Ting L., Urbain T., Komatsubara T.,  
 RA Hatanoto T., Okamoto M., Tojo H.;  
 RT "Identification of functional domains of rat intestinal phospholipase  
 B/lipase. Its cDNA cloning, expression, and tissue distribution.";  
 RL J. Biol. Chem. 273:2222-2231(1998).  
 DR EMBL; D63648; BAA23813.1; -.  
 DR InterPro: IPR001087; Lipase\_GDSL.  
 DR Pfam: PF00657; Lipase\_GDSL; 3.  
 DR PROSITE: PS01098; LIPASE\_GDSL\_SER; 2.  
 SQ SEQUENCE 1450 AA; 161088 MW; 4555898C8FD91F45 CRC64;

Query Match 63.8%; Score 37; DB 11; Length 1450;  
 Best Local Similarity 41.7%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 4; Mismatches 3; Indels 0; Gaps 0;  
 QY 1 YXXRGIRPVGR 12  
 Db 353 YQARQLKPIGRF 364

RESULT 7  
 ID 097950 PRELIMINARY; PRT; 426 AA.  
 AC 097950;  
 DT 01-OCT-2001 (TREMblrel. 18, Created)  
 DT 01-OCT-2001 (TREMblrel. 18, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE TVG1120468 PROTEIN.  
 GN TVG1120468.  
 OS Thermoplasma volcanum.  
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;  
 OC Thermoplasma.  
 OX NCBI\_TaxID=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=GSI / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima T., Ohya Y., Matanabe K., Yamazaki M., Kanehori K., Kawamoto T.,  
 RA Nunoshima T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 sequence of Thermoplasma volcanum.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 DR EMBL; AP000995; BAB60232.1; -.  
 DR InterPro: IPR002881; DUF58.  
 DR InterPro: IPR002035; VWFA.  
 DR Pfam: PF01882; DUF58; 1.  
 DR SMART; SM00327; VWFA; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 426 AA; 49092 MW; 63967CF711116FA4 CRC64;

Query Match 62.1%; Score 36; DB 17; Length 426;  
 Best Local Similarity 75.0%; Pred. No. 61;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 5 GIRPVGR 12  
 Db 304 GIRPVGR 311

RESULT 8  
 ID 09A7W7 PRELIMINARY; PRT; 545 AA.  
 AC 09A7W7;  
 DT 01-JUN-2001 (TREMblrel. 17, Created)  
 DT 01-JUN-2001 (TREMblrel. 17, last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
 DE HYPOTHETICAL PROTEIN CC1602.  
 GN CC1602.  
 OS Caulobacter crescentus.  
 OC Bacteria; Proteobacteria; alpha subdivision; Caulobacter group;  
 OC Caulobacter.  
 OX NCBI\_TaxID=69394;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19089 / CB15;  
 RX MEDLINE=21173698; PubMed=11259647;  
 RA Nielsen W.C., Feldblyum T.V., Laub M.T., Paulsen I.T., Nelson K.E.,  
 RA Eilen J., Heidelberg J.F., Alley M.R.K., Ohta N., Maddock J.R.,  
 RA Pockock J., Nelson W.C., Newton W.C., Stephens C., Phadke N.D., Ely B.,  
 RA DebRoy R.T., Dodson R.J., Durkin A.S., Gilm M.L., Helt D.H.,  
 RA Kolonay J.F., Selt J., Craven M.B., Khouri H., Shetty J., Berry K.,  
 RA Uterback T., Tran K., Wolf A., Vamathevan J., Etmoleva M., White O.,  
 RA Salzberg S.L., Venter J.C., Shapiro L., Fraser C.M.;  
 RT "Complete genome sequence of Caulobacter crescentus.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:4136-4141(2001).  
 DR EMBL; AE005835; AAK23581.1; -.  
 DR TIGR; CC1602; -.  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 545 AA; 60175 MW; FB3187D64F44C238 CRC64;

Query Match 62.1%; Score 36; DB 16; Length 545;  
 Best Local Similarity 50.0%; Pred. No. 80;  
 Matches 6; Conservative 2; Mismatches 4; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
 | : : : : :  
 DB 245 YRAGVPRVGR 256

RESULT 9  
 ID 050431 PRELIMINARY; PRT; 674 AA.  
 AC 050431;  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE 2,4-DIENOLYL-COA REDUCTASE.  
 GN FADH OR RV1175C OR MTV005.11C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 Davies K., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,  
 Horvath T., Jorgensen K., Krogan A., Molese J., Moule S., Murphy L.,  
 Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 Ruster S., Seeger K., Skellern S., Squares R., Squares R.,  
 Sutton J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RA Sultson J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 DR EMBL: AL010186; CNA15852.1; -  
 DR Tuberclist; Rv1175c; -  
 DR InterPro: IPR001064; Crystallin.  
 DR InterPro: IPR001327; FAD\_pyr\_redox.  
 DR InterPro: IPR000205; NAD\_binding.  
 DR InterPro: IPR001155; Oxidored\_FMN.  
 DR Pfam: PF00724; oxidored\_FMN; 1.  
 DR Pfam: PF00070; pyr\_redox; 1.  
 DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
 KW Complete proteome.  
 SQ SEQUENCE 674 AA; 72901 MW; 8f478f74ED3E0A5 CRC64;

Query Match 62.1%; Score 36; DB 16; Length 674;  
 Best Local Similarity 85.7%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 5 GRRPVR 11  
 | : : : : :  
 DB 492 GRRPVR 498

RESULT 10  
 ID 09M371 PRELIMINARY; PRT; 790 AA.  
 AC 09M371;  
 DT 01-OCT-2000 (TREMBLrel. 15, Created)  
 DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HYPOTHETICAL 87.4 KDA PROTEIN.  
 GN F15G16.60.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;

OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA De Haan M., Mearse A.C., Grivell L.A., Mewes H.W., Lemcke K.,  
 Mayer K.F.X., Queller F., Salanoubat M.;  
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL132959; CAB71097.1; -  
 KW Hypothetical protein.  
 SQ SEQUENCE 790 AA; 87376 MW; B222724B75690F30 CRC64;

Query Match 62.1%; Score 36; DB 10; Length 790;  
 Best Local Similarity 50.0%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 3; Mismatches 3; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGRF 12  
 | : : : : :  
 DB 373 YGSRGIRPVGR 384

RESULT 11  
 ID 096MP2 PRELIMINARY; PRT; 144 AA.  
 AC 096MP2;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CDNA FLJ32096 F1S, CLONE OCBF2001075.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN;  
 RA Niinomiya K., Wagatsuna M., Kanda K., Kondo H., Yokoi T., Kodaira H.,  
 Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K., Kamihara K.,  
 Katsuta N., Sato K., Tanikawa M., Yamazaki M., Sugiyama T., Irie R.,  
 Otsuki T., Sato H., Ota T., Wakamatsu A., Ishii S., Yamamoto J.,  
 Isono Y., Kawai-Hiro Y., Saito K., Nishikawa T., Kinura K.,  
 Yamashita H., Matsuo K., Nakamura Y., Sekine M., Kikuchi H.,  
 Murakawa K., Kanehori K., Takahashi-Fujii A., Oshima A., Sugiyama A.,  
 Kawakami B., Suzuki Y., Sugano S., Nagahara K., Masuno Y., Nagai K.,  
 Isogai T.;  
 RT "NEO human cDNA sequencing project."  
 RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK056658; BAB71243.1; -  
 SQ SEQUENCE 144 AA; 15589 MW; 9CCD9EAE7F735881 CRC64;

Query Match 60.3%; Score 35; DB 4; Length 144;  
 Best Local Similarity 60.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 3; Indels 0; Gaps 0;

OY 1 YXXRGIRPVGR 10  
 | : : : : :  
 DB 87 YAGRGIRPVGR 96

RESULT 12  
 ID 097SU7 PRELIMINARY; PRT; 178 AA.  
 AC 097SU7;  
 DT 01-OCT-2001 (TREMBLrel. 18, Created)  
 DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE RIBOSOMAL PROTEIN L6.  
 GN SP0225.



OS Streptococcus pneumoniae.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;  
OC Streptococcus.  
OX NCBI\_TaxID=1313;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TIGR4;  
RX MEDLINE=21357209; PubMed=11463916;  
RA Tettein H., Nelson K.E., Paulsen I.T., Eisen J.A., Read T.D.,  
Peterson S., Heidelberg J., Deboy R.T., Haft D.H., Dodson R.J.,  
Durkin A.S., Gwinn M., Kolonay J.F., Nelson W.C., Peterson J.D.,  
Umayam L.A., White O., Salzberg S.L., Lewis M.R., Radune D.,  
Holtapple E., Khouri H., Wolf A.M., Utterback T.R., Hansen C.L.,  
McDonald L.A., Feldblyum T.V., Angiuoli S., Dickinson T., Hickey E.K.,  
Holt I.E., Loftus B.J., Yang F., Smith H.O., Venter J.C.,  
R.A. Dougherty B.A., Morrison D.A., Hollingshead S.K., Fraser C.M.;  
RT "Complete genome sequence of a virulent isolate of Streptococcus  
pneumoniae.";  
RT Science 293:498-506(2001).  
RL EMBL: AE007336; AAK74405.1; -.  
DR TIGR: SP0225;  
DR InterPro: IPR000702; Ribosomal\_L6.  
DR InterPro: IPR002358; Ribosomal\_L6\_1.  
DR Pfam: PF00347; Ribosomal\_L6; 1.  
DR PRINTS: PR00059; RIBOSOMAL\_L6; 1.  
DR PRODOM: PD002236; Ribosomal\_L6; 1.  
DR PROSITE: PS00525; RIBOSOMAL\_L6\_1; 1.  
KM Complete proteome.  
SQ SEQUENCE 178 AA; 19431 MW; D2C99A157C01F13D CRC64;

Query Match 60.3%; Score 35; DB 16; Length 178;  
Best Local Similarity 58.3%; Pred. No. 37;  
Matches 7; Conservative 1; Mismatches 4; Indels 0; Gaps 0;  
QY 1 YXKGIKPVGRF 12  
DB 157 YKKGIRYGEF 168  
RESULT 13  
042924  
ID 042924 PRELIMINARY; PRT; 272 AA.  
AC 042924;  
DT 01-JAN-1999 (TREMBLrel. 09, Created)  
DT 01-JAN-1999 (TREMBLrel. 09, Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 30.9 KDA PROTEIN C16A3.19 IN CHROMOSOME II.  
GN SPBC16A3.19.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;  
RT Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL021748; CAIL6870.1; -.  
DR InterPro: IPR001005; MYD\_DNA\_bind.  
DR PROSITE: PS00037; MYB\_1; UNKNOWNW\_1.  
KM Hypothetical protein.  
SQ SEQUENCE 272 AA; 30901 MW; 8FE0EB29217107A8 CRC64;

Query Match 60.3%; Score 35; DB 3; Length 272;  
Best Local Similarity 85.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 4 RGIRPVG 10  
DB 47 RGLRPVG 53

RESULT 14  
0981V3  
ID 0981V3 PRELIMINARY; PRT; 371 AA.  
AC 0981V3;  
DT 01-OCT-2001 (TREMBLrel. 18, Created)  
DT 01-OCT-2001 (TREMBLrel. 18, Last sequence update)  
DE MR9222-2001 (TREMBLrel. 18, Last annotation update)  
GN MR9222.  
OS Rhizobium loti (Mesorhizobium loti).  
OG Plasmid pMla.  
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
OC Rhizobiaceae; Mesorhizobium.  
OX NCBI\_TaxID=381;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MAF30309;  
RX MEDLINE=21082930; PubMed=11214968;  
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,  
Matsunaga T., Nakamura K., Ishikawa A., Kawashima K., Kimura T.,  
Kishida Y., Kiyokawa C., Kohara M., Matsunoto M., Matsuno A.,  
Mochizuki Y., Nakayama S., Nakazaki N., Shimpo S., Sugimoto M.,  
Takeuchi C., Yamada M., Tabata S.;  
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium  
Mesorhizobium loti.";  
RT DNA Res. 7:331-338(2000).  
RL EMBL: AF003015; BAB54606.1; -.  
DR Plasmid; Complete proteome.  
KM SEQUENCE 371 AA; 42385 MW; 98D9874B5E7831D1 CRC64;  
SQ

Query Match 60.3%; Score 35; DB 16; Length 371;  
Best Local Similarity 62.5%; Pred. No. 83;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
QY 4 RGIRPVGR 11  
DB 179 RGVPIGR 186  
RESULT 15  
053895  
ID 053895 PRELIMINARY; PRT; 504 AA.  
AC 053895;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE PUTATIVE TWO-COMPONENT SENSOR.  
GN RV0982 OR MTW044.10.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37R;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Broesch R., Parthill J., Garner T., Churcher C., Harris D.,  
Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekala F.,  
Radocek K., Basham D., Brown D., Chillingworth T., Connor R.,  
Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
Hornaby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence.";  
RT Nature 393:537-544(1998).  
RL -1- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.  
DR EMBL: AL021999; CAIL17581.1; -.  
DR HSSP; P02933; IBXD.

DR Tuberculin: Rv0982; -  
 DR InterPro: IPR003660; HAM.  
 DR InterPro: IPR003594; HATPase\_C.  
 DR InterPro: IPR003661; His\_kinA.  
 DR InterPro: IPR004359; His\_kinA.  
 DR Pfam: PF00672; HAM; 1.  
 DR Pfam: PF02518; HATPase\_C; 1.  
 DR Pfam: PF00512; signal; 1.  
 DR SMART; SM00304; HAM; 1.  
 DR SMART; SM00387; HATPase\_C; 1.  
 DR SMART; SM00388; HisKA; 1.  
 KW Complete proteome: Kinase; Phosphorylation; Sensory transduction;  
 KW Transferase.  
 SQ SEQUENCE 504 AA; 54407 MM; 26792A71AC432232 CRC64;

Query Match 60.3%; Score 35; DB 16; Length 504;  
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
 QY 5 GIRPVGR 11  
 1:|||||  
 DB 188 GLRPVGR 194

Search completed: September 13, 2002, 09:29:23  
 Job time: 1068 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:35 ; Search time 399.68 Seconds  
(without alignments)  
3.335 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_10\_21

Perfect score: 51

Sequence: 1 XXRGIRPVGRFX 12

## Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

## Database :

A\_Geneseq\_032802:.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneeqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	94.1	10	20	AAW95178
2	48	94.1	15	18	AAW91400
3	48	94.1	15	20	AAW97230
4	48	94.1	15	21	AAV49283
5	48	94.1	15	21	AAV49295
6	48	94.1	15	21	AAV49296
7	48	94.1	20	18	AAW41394
8	48	94.1	20	18	AAW41387
9	48	94.1	20	18	AAW41374
10	48	94.1	20	20	AAW97232
11	48	94.1	20	20	AAW97234

12	48	94.1	20	20	AAW97236	Human type ligand
13	48	94.1	20	20	AAW95191	Bovine pituitary-d
14	48	94.1	20	20	AAW95175	Murine pituitary-d
15	48	94.1	20	21	AAW10350	Bovine oxytocin se
16	48	94.1	20	21	AAW10358	Rat oxytocin secre
17	48	94.1	20	21	AAW10365	Human oxytocin sec
18	48	94.1	20	21	AAW10369	Oxytocin secretion
19	48	94.1	20	21	AAV49294	19P2 ligand peptid
20	48	94.1	20	21	AAV49301	19P2 ligand peptid
21	48	94.1	20	21	AAV49302	19P2 ligand peptid
22	48	94.1	20	22	AAW62519	Bovine CRH releas
23	48	94.1	20	22	AAW62527	Rat CRH releas
24	48	94.1	20	22	AAW62534	Human CRH releas
25	48	94.1	20	22	AAW62538	CRH releasing prot
26	48	94.1	20	22	AAW62538	Prolactin releas
27	48	94.1	20	22	AAW62538	Prolactin releas
28	48	94.1	20	22	AAW62538	Prolactin releas
29	48	94.1	20	22	AAW62538	Prolactin releas
30	48	94.1	20	22	AAW62538	Prolactin releas
31	48	94.1	20	22	AAW62538	Prolactin releas
32	48	94.1	20	22	AAW62538	Prolactin releas
33	48	94.1	20	22	AAW62538	Prolactin releas
34	48	94.1	20	22	AAW62538	Prolactin releas
35	48	94.1	20	22	AAW62538	Prolactin releas
36	48	94.1	20	22	AAW62538	Prolactin releas
37	48	94.1	20	22	AAW62538	Prolactin releas
38	48	94.1	20	22	AAW62538	Prolactin releas
39	48	94.1	20	22	AAW62538	Prolactin releas
40	48	94.1	20	22	AAW62538	Prolactin releas
41	48	94.1	20	22	AAW62538	Prolactin releas
42	48	94.1	20	22	AAW62538	Prolactin releas
43	48	94.1	20	22	AAW62538	Prolactin releas
44	48	94.1	20	22	AAW62538	Prolactin releas
45	48	94.1	20	22	AAW62538	Prolactin releas

## ALIGNMENTS

RESULT 1	
AAW95178	standard; Protein: 10 AA.
AAW95178	
10-MAR-1999	(first entry)
DE	Murine pituitary-derived ligand polypeptide antigenic epitope.
KW	Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
KW	ltsue; screen; therapeutic; binding; senile dementia; ligand; murine;
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
KW	Creatin; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
KW	gene therapy; transgenic animal; epitope.
XX	
OS	Mus sp.
PN	W09849295-A1.
PD	05-NOV-1998.
PE	27-APR-1998; 98WO-JP01923.
PR	28-APR-1997; 97JP-0109974.
PA	(TAKE ) TAKEDA CHEM IND LTD.
PI	Fukushima S, Hinuma S;
DR	WPI: 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening  
PS Disclosure; Page 26; 206pp; English.  
XX  
CC The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or GHR-1 (rat). Cells transfected with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutant are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.  
XX  
SQ Sequence 10 AA;  
  
Query Match 94.1%; Score 48; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 RGIRPVGRF 11  
| | | | | | | | | |  
Db 2 rglrtpvgrf 10  
  
RESULT 2  
AAW31400  
ID AAW31400 standard; Peptide; 15 AA.  
XX  
AC AAW31400;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Synthetic ligand 19P2-L31 peptide II.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent; antigen.  
XX  
OS Synthetic.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
FUJII R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;

XX  
DR WPI; 1997-363672/33.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Example 43; Page 151; 258pp; English.  
XX  
CC This peptide contains the partial C-terminal sequence of the synthetic  
CC ligand polypeptide 19P2-L31 which is capable of binding to a G  
CC protein-coupled receptor protein. This peptide is used as an antigen to  
CC prepare rabbit anti-bovine 19P2-L31 antibodies which are used in binding  
CC assays. Pharmaceutical compositions containing this ligand may be used  
CC as a pituitary function modulator, a central nervous system modulator  
CC or a pancreatic function modulator. This ligand could have specific  
CC applications as a prophylactic or therapeutic agent for dementia,  
CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
CC hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia,  
CC hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
CC and/or oligogalactia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand  
CC affecting activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 15 AA;  
  
Query Match 94.1%; Score 48; DB 18; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
  
QY 3 RGIRPVGRF 11  
| | | | | | | | | |  
Db 7 rglrtpvgrf 15  
  
RESULT 3  
AAW97230  
ID AAW97230 standard; Peptide; 15 AA.  
XX  
AC AAW97230;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE C-terminal ligand polypeptide derived antigen.  
XX  
KW G protein-coupled receptor; GPCR; hypovarianism; gonecyst caecogenesis;  
KW menopausal syndrome; eutthyroid; hypometabolism; lactation; modulation;  
KW pituitary adenomatosis; brain tumour; emmenopathy; autolimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriochorionoma; hydatid mole;  
KW interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia; prolactin secretion.  
XX  
OS Synthetic.  
XX  
PN WO9858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
FUJII R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI; 1999-105614/09.  
XX  
XX Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
XX Example 43; Page 123; 241pp; English.  
XX  
XX AAM97229-31 represent a ligand polypeptide derived fragments used to  
CC produce antibodies. The specification describes an agent for modulating  
CC prolactin secretion which comprises a ligand polypeptide or a salt, for  
CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
CC prolactin secretion can be used for treating or preventing  
CC hypogonadism, gonocyst cacosgenesis, menopausal syndrome, euthyroid or  
CC hypometabolism. They can be used for promoting lactation in a domestic  
CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
CC secretion can be used for treating or preventing pituitary adenomatosis,  
CC brain tumour, emmeniopathy, autoimmune disease, prolactinoma,  
CC infertility, impotence, amenorrhea, galactorrhea, acromegaly,  
CC Chlari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright  
CC syndrome, lymphoma, Sheehan syndrome or dysospermia. The inhibitory  
CC agents can also be used as contraceptives. The agents for modulating  
CC placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, irruption mole, abortion, unthrifty  
CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
CC oxytocia.  
XX  
XX Sequence 15 AA:  
SQ  
Query Match 94.1%; Score 48; DB 20; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 RGIRPVGRF 11  
Db 7 rglrpygrf 15  
RESULT 4  
AAV49293  
ID AAV49293 standard; peptide: 15 AA.  
XX  
XX AAV49293;  
AC  
XX  
XX 22-FEB-2000 (first entry)  
DT  
XX  
XX 19P2 ligand peptide fragment.  
DE  
XX  
XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
KM pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FT Modified-site 15  
FT /note= "C-terminal amide"  
XX  
XX WO9960112-A1.  
PN  
XX  
XX 25-NOV-1999.  
PD  
XX  
XX 20-MAY-1999; 99WO-JP02650.  
PF  
XX  
XX 21-MAY-1998; 98JP-0140293.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
PI  
XX  
XX WPI; 2000-039381/03.  
DR  
XX

PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
XX Disclosure; Page 26; 73pp; Japanese.  
XX  
XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
XX Sequences AAV49290-302 represent peptide fragments of the 19P2 ligand.  
XX  
XX Sequence 15 AA:  
SQ  
Query Match 94.1%; Score 48; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 RGIRPVGRF 11  
Db 7 rglrpygrf 15  
RESULT 5  
AAV49295  
ID AAV49295 standard; peptide: 15 AA.  
XX  
XX AAV49295;  
AC  
XX  
XX 22-FEB-2000 (first entry)  
DT  
XX  
XX 19P2 ligand peptide fragment.  
DE  
XX  
XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
KM pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FT Modified-site 1  
FT /note= "N-terminal acetylation"  
FT Modified-site 4  
FT /note= "acetylated Tyr"  
FT Modified-site 15  
FT /note= "C-terminal amide"  
XX  
XX WO9960112-A1.  
PN  
XX  
XX 25-NOV-1999.  
PD  
XX  
XX 20-MAY-1999; 99WO-JP02650.  
PF  
XX  
XX 21-MAY-1998; 98JP-0140293.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
PI  
XX  
XX WPI; 2000-039381/03.  
DR  
XX  
XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
XX Disclosure; Page 26; 73pp; Japanese.  
XX  
XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function



Db 12 rgirpvgf 20

## RESULT 8

AAW31387  
ID AAW31387 standard; Peptide; 20 AA.

AC AAW31387;

DT 06-APR-1998 (first entry)

DE Rat type G protein-coupled receptor ligand fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.

XX Rat sp.

PN WO9724436-A2.

PD 10-JUL-1997.

PE 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

DR WPI; 1997-363672/33.

DR N-PSDB; AAV02424.

XX Ligand peptide for G protein-coupled receptor - acts by modulating

PT function in the central nervous system; pancreas and pituitary gland

XX Claim 2; Page 180; 258pp; English.

XX This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the  
CC sequence represented in AAW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hypercholesterolemia, hyperglycemia, hyperlipidemia,  
CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 94.1%; Score 48; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
GY 3 RGIRPVGRF 11  
|||||

Db 12 rgirpvgf 20

## RESULT 9

AAW31374  
ID AAW31374 standard; Peptide; 20 AA.

AC AAW31374;

DT 06-APR-1998 (first entry)

DE Bovine G protein-coupled receptor ligand peptide fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.

XX Bos taurus.

PN WO9724436-A2.

PD 10-JUL-1997.

PE 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

DR WPI; 1997-363672/33.

DR N-PSDB; AAV02397.

XX Ligand peptide for G protein-coupled receptor - acts by modulating

PT function in the central nervous system; pancreas and pituitary gland

XX Claim 2; Page 161; 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary  
CC derived ligand corresponding to amino acid residues 34 to 53 of the  
CC sequence in AAW31368 and is used in an assay to monitor ligand binding  
CC to the G protein-coupled receptor protein. Pharmaceutical compositions  
CC containing this ligand may be used as a pituitary function modulator, a  
CC central nervous system modulator or a pancreatic function modulator.  
CC This ligand could have specific applications as a prophylactic or  
CC therapeutic agent for dementia, depression, hyperkinetic syndrome,  
CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
CC growth hormone secretory disease, hyper- and polyphagia,  
CC hypercholesterolemia, hyperglycemia, hyperlipidemia,  
CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischemia, epilepsy, amyotrophic lateral sclerosis, acute  
CC myocardial infarction, spinocerebellar degeneration, bone fracture,  
CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility  
CC and/or oligogalactia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand thus  
CC affecting activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 94.1%; Score 48; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
GY 3 RGIRPVGRF 11  
|||||

```
RESULT 10
AAW97232
ID AAW97232 standard; peptide; 20 AA.
XX
AC AAW97232;
XX
DT 06-MAY-1999 (first entry)
XX
DE Bovine pituitary-derived ligand polypeptide fragment.
XX
KW Bovine pituitary-derived ligand; modulation; prolactin secretion;
KW G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacogenesis;
KW menopausal syndrome; eutthyroid; hypometabolism; lactation;
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
KW contraceptive; placental function; choriocarcinoma; hydatid mole;
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
KW abnormal lipidmetabolism; oxytocia.
XX
OS Bos sp.
XX
PN WO9858962-A1.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-JP02765.
XX
PR 23-JUN-1997; 97JP-0165437.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;
XX
DR WPI: 1999-105614/09.
XX
PT Use of G protein-coupled receptor ligands - for modulating prolactin
PT secretion or placental function, e.g. for treating menopausal
PT syndrome, tumours, autoimmune disease or abnormal pregnancy
XX
PS Claim 3; Page 136; 241pp; English.
XX
CC The present sequence represents a bovine pituitary-derived ligand
CC fragment. It is used in the course of the invention. The specification
CC describes an agent for modulating prolactin secretion which comprises a
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
CC protein. The agents for promoting prolactin secretion can be used for
CC treating or preventing hypovarianism, gonecyst cacogenesis, menopausal
CC syndrome, eutthyroid or hypometabolism. They can be used for promoting
CC lactation in a domestic mammal and as an aphrodisiac. The agents for
CC inhibiting prolactin secretion can be used for treating or preventing
CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
CC The inhibitory agents can also be used as contraceptives. The agents for
CC modulating placental function can be used for treating or preventing
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.
XX
SQ Sequence 20 AA:

Query Match 94.1%; Score 48; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIPIVGRF 11
|||||
DB 12 rglrpygrf 20
```

```
RESULT 11
AAW97234
ID AAW97234 standard; peptide; 20 AA.
XX
AC AAW97234;
XX
DT 06-MAY-1999 (first entry)
XX
DE Rat type ligand polypeptide fragment.
XX
KW Rat type ligand; modulation; prolactin secretion;
KW G protein-coupled receptor; GPCR; hypovarianism; gonecyst cacogenesis;
KW menopausal syndrome; eutthyroid; hypometabolism; lactation;
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;
KW contraceptive; placental function; choriocarcinoma; hydatid mole;
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;
KW abnormal lipidmetabolism; oxytocia.
XX
OS Rattus sp.
XX
PN WO9858962-A1.
XX
PD 30-DEC-1998.
XX
PF 22-JUN-1998; 98WO-JP02765.
XX
PR 23-JUN-1997; 97JP-0165437.
XX
PA (TAKE ) TAKEDA CHEM IND LTD.
XX
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;
XX
DR WPI: 1999-105614/09.
XX
PT Use of G protein-coupled receptor ligands - for modulating prolactin
PT secretion or placental function, e.g. for treating menopausal
PT syndrome, tumours, autoimmune disease or abnormal pregnancy
XX
PS Claim 3; Page 154; 241pp; English.
XX
CC The present sequence represents a rat type ligand fragment. It
CC is used in the course of the invention. The specification describes
CC an agent for modulating prolactin secretion which comprises a
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)
CC protein. The agents for promoting prolactin secretion can be used for
CC treating or preventing hypovarianism, gonecyst cacogenesis, menopausal
CC syndrome, eutthyroid or hypometabolism. They can be used for promoting
CC lactation in a domestic mammal and as an aphrodisiac. The agents for
CC inhibiting prolactin secretion can be used for treating or preventing
CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.
CC The inhibitory agents can also be used as contraceptives. The agents for
CC modulating placental function can be used for treating or preventing
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.
XX
SQ Sequence 20 AA:

Query Match 94.1%; Score 48; DB 20; Length 20;
Best Local Similarity 100.0%; Pred. No. 0.011;
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIPIVGRF 11
|||||
DB 12 rglrpygrf 20
```



RESULT 12  
 AAM97236  
 ID AAM97236 standard; peptide: 20 AA.  
 XX  
 AC AAM97236.  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE Human type ligand polypeptide fragment.  
 XX  
 KM Rat type ligand: modulation: prolactin secretion;  
 KM G protein-coupled receptor: GPCR; hypovarianism; gonocyst cacosgenesis;  
 KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KM pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KM contraceptive; placental function; choriorcarcinoma; hydallid mole;  
 KM interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KM abnormal lipidmetabolism; oxytocia.  
 XX  
 OS Homo sapiens.  
 XX  
 PN WO9858962-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 22-JUN-1998: 98WO-JP02765.  
 XX  
 PR 23-JUN-1997: 97JP-0165437.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H:  
 DR WPI: 1999-105614/09.  
 XX  
 PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Claim 3: Page 166; 241pp; English.  
 XX  
 CC The present sequence represents a human type ligand fragment. It  
 CC is used in the course of the invention. The specification describes  
 CC an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenopathy, galactorrhea,  
 CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing for  
 CC choriorcarcinoma, hydallid mole, interruption mole, abortion, unthrifty fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 CC  
 XX  
 SQ Sequence 20 AA;  
 Query Match 94.1%; Score 48; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 3 RGIRPVGRF 11  
 DB 12 rgirpvgrf 20

RESULT 13  
 AAM95191  
 ID AAM95191 standard; peptide: 20 AA.  
 XX  
 AC AAM95191.  
 XX  
 DT 10-MAR-1999 (first entry)  
 XX  
 DE Bovine pituitary-derived ligand polypeptide fragment.  
 XX  
 KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
 KM GPR10; UHR-L; modulator; pituitary; central nervous system; pancreas;  
 KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KM gene therapy; transgenic animal; bovine.  
 XX  
 OS Bos sp.  
 XX  
 PN WO9849295-A1.  
 XX  
 PD 05-NOV-1998.  
 XX  
 PF 27-APR-1998: 98WO-JP01923.  
 XX  
 PR 28-APR-1997: 97JP-0109974.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fukusumi S, Hinuma S:  
 DR WPI: 1999-009423/01.  
 XX  
 PT New polypeptide ligand for orphan G protein coupled receptors - used  
 PT for treating disorders of central nervous system, pituitary and  
 PT pancreas, and for drug screening  
 XX  
 PS Example 19: Page 151; 206pp; English.  
 XX  
 CC The invention relates to a murine pituitary-derived ligand polypeptide  
 CC which is a ligand for the G-protein coupled orphan receptor designated  
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
 CC the ligand polypeptide encoding DNA are used to produce a recombinant  
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
 CC function of the pituitary, central nervous system, pancreas and other  
 CC tissues and can be used to screen for agents that modulate binding of the  
 CC polypeptide to the receptor; to quantify the amount of receptor in a  
 CC sample and to raise antibodies. They may also be used therapeutically,  
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
 CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;  
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
 CC operative nutritional status and as vasopressor. Transgenic animals  
 CC carrying the ligand polypeptide encoding DNA or its mutant are used to  
 CC study the function of the polypeptide-expressing genes, as models of  
 CC disease, for drug screening and as source of cell lines. The ligand  
 CC polypeptide DNA is used as a source of probes and primers; to identify  
 CC related sequences; in receptor-binding assays; for production of Ab and  
 CC antisera; in drug development; for gene therapy and to develop transgenic  
 CC animals. The present sequence represents a bovine genome-derived ligand  
 CC polypeptide fragment which is similar to the murine ligand-polypeptide.  
 CC  
 XX  
 SQ Sequence 20 AA;  
 Query Match 94.1%; Score 48; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 3 RGIRPVGRF 11

Db 12 rglrpygrf 20

RESULT 14  
AAW95175  
ID AAW95175 standard: Protein: 20 AA.

AC AAW95175;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; epitope.

OS Mus sp.

PN WO9849295-A1.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

DR WPI: 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening  
PS Disclosure; Page 26: 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.

SQ Sequence 20 AA;

Query Match 94.1%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 3 RGI RPYGRF 11  
Db 12 rglrpygrf 20

RESULT 15  
AAB10350  
ID AAB10350 standard: peptide; 20 AA.

AC AAB10350;

DT 24-NOV-2000 (first entry)

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 6.

KM Bovine: oxytocin secretion promoter; G protein-coupled receptor protein;  
KM treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
KM caesarean section; artificial fertilization; galactostasis; goat; pig;  
KM veterinary medicine; milk production.

OS Bos taurus.

PN WO200038704-A1.

PD 06-JUL-2000.

PF 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0365585.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI: 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G  
PT protein-coupled receptor protein, for promoting secretion of oxytocin,  
PT as drugs for diseases relating to oxytocin secretion and in veterinary  
PT medicine -  
PS Claim 5; Page 51; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent  
CC which contains a ligand peptide or its salt for the G protein-coupled  
CC receptor protein. It is useful in the form of drugs for ameliorating,  
CC preventing and treating diseases relating to oxytocin secretion e.g.  
CC weak pains and atonic bleeding, before and after expulsion of placenta,  
CC uterine recovery failure, caesarean section, stoppage of artificial  
CC fertilization or galactostasis and is also applicable in veterinary  
CC medicine for promoting milk production in cow, goat and pig. This  
CC sequence represents a bovine peptide which acts as an oxytocin secretion  
CC promoter.

SQ Sequence 20 AA;

Query Match 94.1%; Score 48; DB 21; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGI RPYGRF 11  
Db 12 rglrpygrf 20

Search completed: September 13, 2002, 09:18:36  
Job time: 501 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 Seconds  
(without alignments)  
2.145 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_10\_21  
Perfect score: 51  
Sequence: 1 XXRGIRPYGRFX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5h\_COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5b\_COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6a\_COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6b\_COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCTUS\_COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfills1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	94.1	15	4	US-08-776-971-93
2	48	94.1	20	3	US-09-105-678A-34
3	48	94.1	20	3	US-09-105-678A-40
4	48	94.1	20	3	US-09-105-678A-46
5	48	94.1	20	4	US-08-776-971-8
6	48	94.1	20	4	US-08-776-971-50
7	48	94.1	20	4	US-08-776-971-64
8	48	94.1	20	4	US-08-776-971-98
9	48	94.1	20	4	US-09-421-208-34
10	48	94.1	20	4	US-09-421-208-40
11	48	94.1	20	4	US-09-421-208-46
12	48	94.1	21	3	US-09-105-678A-28
13	48	94.1	21	3	US-09-105-678A-35
14	48	94.1	21	3	US-09-105-678A-41
15	48	94.1	21	3	US-09-105-678A-47
16	48	94.1	21	4	US-08-776-971-9
17	48	94.1	21	4	US-08-776-971-51
18	48	94.1	21	4	US-08-776-971-65
19	48	94.1	21	4	US-08-776-971-28
20	48	94.1	21	4	US-09-421-208-35
21	48	94.1	21	4	US-09-421-208-41
22	48	94.1	21	4	US-09-421-208-47
23	48	94.1	22	3	US-09-105-678A-36
24	48	94.1	22	3	US-09-105-678A-42
25	48	94.1	22	3	US-09-105-678A-48
26	48	94.1	22	4	US-08-776-971-10
27	48	94.1	22	4	US-08-776-971-52

28	48	94.1	22	4	US-08-776-971-66	Sequence 66, Appl
29	48	94.1	22	4	US-08-776-971-73	Sequence 73, Appl
30	48	94.1	22	4	US-09-421-208-36	Sequence 36, Appl
31	48	94.1	22	4	US-09-421-208-42	Sequence 42, Appl
32	48	94.1	22	4	US-09-421-208-48	Sequence 48, Appl
33	48	94.1	31	3	US-09-105-678A-7	Sequence 7, Appl1
34	48	94.1	31	3	US-09-105-678A-8	Sequence 8, Appl1
35	48	94.1	31	3	US-09-105-678A-9	Sequence 9, Appl1
36	48	94.1	31	3	US-09-105-678A-31	Sequence 31, Appl
37	48	94.1	31	3	US-09-105-678A-37	Sequence 37, Appl
38	48	94.1	31	3	US-09-105-678A-43	Sequence 43, Appl
39	48	94.1	31	4	US-09-172-353-4	Sequence 4, Appl1
40	48	94.1	31	4	US-08-776-971-5	Sequence 5, Appl1
41	48	94.1	31	4	US-08-776-971-47	Sequence 47, Appl
42	48	94.1	31	4	US-08-776-971-61	Sequence 61, Appl
43	48	94.1	31	4	US-08-776-971-97	Sequence 97, Appl
44	48	94.1	31	4	US-09-421-208-7	Sequence 7, Appl1
45	48	94.1	31	4	US-09-421-208-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-08-776-971-93  
; Sequence 93, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Hada, Yugo  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kikada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27, 026  
; REFERENCE/DOCKET NUMBER: 47116  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-08-776-971-93

Query Match 94.1%; Score 48; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0024;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
|||||  
DB 7 RGIRPYGRF 15

## RESULT 2

US-09-105-678A-34  
Sequence 34, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 94.1%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
|||||  
DB 12 RGIRPYGRF 20

RESULT 3  
US-09-105-678A-40  
Sequence 40, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 94.1%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
|||||  
DB 12 RGIRPYGRF 20

## RESULT 4

US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 94.1%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
Db 12 RGIRPVGRF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
Db 12 RGIRPVGRF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
DB 12 RGIRPVGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
DB 12 RGIRPVGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 98:  
US-08-776-971-98

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPVGRF 11  
12 RGIRPVGRF 20

RESULT 9  
US-09-421-208-34  
; Sequence 34, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPVGRF 11

Db 12 RGIRPVGRF 20

RESULT 10  
US-09-421-208-40  
; Sequence 40, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPVGRF 11  
12 RGIRPVGRF 20

RESULT 11  
US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:

ADDRESSER: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 94.1%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIPIVGRF 11  
Db 12 RGIPIVGRF 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product- "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product- "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product- "Gly-OH or Gly-Arg"  
US-09-105-678A-28

Query Match 94.1%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIPIVGRF 11  
Db 12 RGIPIVGRF 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:



TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 94.1%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||||||  
DB 12 RGIRPVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 94.1%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||||||

DB 12 RGIRPVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 94.1%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||||||  
DB 12 RGIRPVGRF 20

Search completed: September 13, 2002, 09:20:58  
Job time: 623 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:02 ; Search time 172.41 Seconds  
(without alignments)  
6.688 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_10\_21  
Perfect score: 51  
Sequence: 1 XXRGIRPVGRFX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	94.1	83	2 JC7607	prolactin-releasing
2	36	70.6	674	2 G70875	probable oxidoredu
3	35	68.6	272	2 T39533	hypothetical prote
4	35	68.6	391	1 XXG2AC	acetyl-CoA C-acety
5	35	68.6	504	2 B70821	hypothetical prote
6	35	68.6	519	2 T45447	probable two-compo
7	35	68.6	894	2 E82221	DNA gyrase, chain
8	35	68.6	1056	2 B82557	hypothetical prote
9	34	66.7	226	2 A65057	hypothetical prote
10	34	66.7	307	2 S51485	ribosomal protein
11	34	66.7	328	2 S77236	ribosomal protein
12	34	66.7	343	2 AH1823	30S ribosomal prot
13	33	64.7	290	2 F97646	probable aliphatic
14	33	64.7	290	2 AD2870	ABC transporter, m
15	33	64.7	313	2 B95351	Vibrio type IV secr
16	33	64.7	319	2 T36845	probable membrane
17	33	64.7	335	2 E70655	hypothetical prote
18	33	64.7	353	2 C70362	hypothetical prote
19	33	64.7	366	2 C70365	twisting motility
20	33	64.7	445	2 T35893	FAD-dependent oxid
21	33	64.7	477	2 S68367	protoporphyrinogen
22	33	64.7	477	2 S65684	protoporphyrinogen
23	33	64.7	477	2 A56449	protoporphyrinogen
24	33	64.7	485	2 E83663	glutamy-tRNA synt
25	33	64.7	511	2 T40334	hypothetical prote
26	33	64.7	589	2 G87485	hypothetical prote
27	33	64.7	1940	2 F75393	hypothetical prote
28	32	62.7	57	2 I35055	MHC class II histo
29	32	62.7	149	2 AH2262	hypothetical prote

30	32	62.7	150	2 S09872	hypothetical prote
31	32	62.7	154	2 T34825	hypothetical prote
32	32	62.7	249	2 S76104	hypothetical prote
33	32	62.7	266	2 T43770	ribosomal protein
34	32	62.7	285	2 T50932	short-chain dehydr
35	32	62.7	292	2 G95926	probable saccharid
36	32	62.7	299	2 C70643	hypothetical prote
37	32	62.7	307	2 AD0883	conserved hypotet
38	32	62.7	309	2 T29118	hypothetical prote
39	32	62.7	374	2 G70947	hypothetical prote
40	32	62.7	388	2 AG3444	transporter BME115
41	32	62.7	409	2 T51126	hypothetical prote
42	32	62.7	410	2 D83573	tRNA nucleotidyl t
43	32	62.7	435	2 A72658	probable isocitrat
44	32	62.7	492	2 G70749	probable hycE prot
45	32	62.7	562	2 S16594	regulatory protein

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C>Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A>Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; M0ID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 94.1%; Score 48; DB 2; Length 83;  
Best local similarity 100.0%; Pred. No. 0.016;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
Db 44 RGIRPVGRF 52

RESULT 2  
G70875  
probable oxidoreductase (EC 1.5.99.-) - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: G70875  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.  
A>Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; M0ID:9825987  
A:Accession: G70875  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-674 <COL>  
A:Cross-references: GB:AL010186; GB:AL123456; M0D:93261493; PIDN:CAI5852.1; PID:g269  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: fadH

C:Superfamily: Methylphilus methylotrophus W3A1 trimethylamine dehydrogenase  
C:Keywords: 4Fe-4S; metalloprotein; oxidoreductase  
F:337,340,344,356/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 70.6%; Score 36; DB 2; Length 674;  
Best Local Similarity 85.7%; Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 GIRPVR 10  
DB 492 GIRPVR 498

RESULT 3  
195533  
hypothetical protein SPBC16A3.19 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39533  
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
submitted to the EMBL Data Library, February 1998  
A:Reference number: 221862  
A:Accession: T39533  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-272 <WOO>  
A:Cross-references: EMBL:AL021748; PIDN:CAA16870.1; GSPDB:GN00067; SPDB:SPBC16A3.19  
A:Experimental source: strain 972h-; cosmid c16A3  
C:Genetics:  
A:Gene: SPDB:SPBC16A3.19  
A:Map position: 2  
A:Introns: 53/1; 97/3

Query Match 68.6%; Score 35; DB 2; Length 272;  
Best Local Similarity 85.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPV 9  
DB 47 RGIRPV 53

RESULT 4  
XG2AC  
acetyl-CoA C-acetyltransferase (EC 2.3.1.9) - Zoogloea ramigera  
N:Alternate names: acetoacetyl-CoA thiolase; biosynthetic thiolase; thiolase II  
C:Species: Zoogloea ramigera  
C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 05-May-2000  
C:Accession: A26121  
R:Peoples, O.P.; Masamune, S.; Walsh, C.T.; Sinskey, A.J.  
J. Biol. Chem. 262, 97-102, 1987  
A:Title: Biosynthetic thiolase from Zoogloea ramigera. III. Isolation and characterization  
A:Reference number: A26121; MUID:87083504  
A:Accession: A26121  
A:Molecule type: DNA  
A:Residues: 1-391 <PEO>  
A:Cross-references: EMBL:J02631; NID:q155617; PIDN:AAA27706.1; PID:q155618  
A:Experimental source: strain I-16-M, ATCC 19623  
C:Comment: The active enzyme, a tetramer of identical chains, catalyzes the reversible  
hesis, and poly-beta-hydroxybutyrate biosynthesis.  
C:Genetics:  
A:Gene: phbA  
C:Superfamily: acetyl-CoA acetyltransferase  
C:Keywords: acyltransferase; coenzyme A; homotetramer; ketone body metabolism; poly-beta  
F:2-391/Product: acetyl-CoA acetyltransferase #status predicted <MAT>  
F:89/Active site: Cys #status predicted

Query Match 68.6%; Score 35; DB 1; Length 391;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPV 10  
DB 266 RGIRPV 273

RESULT 5  
B70821  
hypothetical protein RV0982 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70821  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Felkewell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:96295987  
A:Accession: B70821  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-504 <COL>  
A:Cross-references: GB:AL021999; GB:AL123456; NID:q3261538; PIDN:CAA17581.1; PID:q291  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV0982  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 68.6%; Score 35; DB 2; Length 504;  
Best Local Similarity 85.7%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 GIRPV 10  
DB 188 GIRPV 194

RESULT 6  
T45447  
probable two-component system sensor [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C>Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T45447  
R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, February 1998  
A:Reference number: 222967  
A:Accession: T45447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-519 <JAM>  
A:Cross-references: EMBL:AL035500; PIDN:CAB36689.1  
A:Experimental source: cosmid L373  
C:Genetics:  
A:Note: MLCB373.27  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 68.6%; Score 35; DB 2; Length 519;  
Best Local Similarity 85.7%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 GIRPV 10  
DB 188 GIRPV 194

RESULT 7  
E82221  
DNA gyrase, chain A VC1258 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C:Accession: E82221  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.J.;  
Chardon, D.; Esmailaei, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, H.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen *Vibrio cholerae*.  
A:Reference number: AB2035; M0ID:20406833  
A:Accession: E82221  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-894 <HEI>  
A:Cross-references: GB:AE004205; GB:AE003852; NID:g9655740; PIDN:AAF94417.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1258  
A:Map position: 1  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

Query Match 68.6%; Score 35; DB 2; Length 894;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 3 GRIIPVGR 10  
|:|:|:|:  
Db 733 KGVPRMGR 740

RESULT 8  
B82557  
hypothetical protein XF2445 [imported] - *Xylella fastidiosa* (strain 9a5c)  
C:Species: *Xylella fastidiosa*  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82557  
R:Anonymous, The *Xylella fastidiosa* Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen *Xylella fastidiosa*.  
A:Reference number: AB2515; M0ID:20365717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1056 <SIM>  
A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF85244.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Britons, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carre, H  
as-Neto, E.; Docena, C.; El-Dor, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laiz  
chido, M.A.; Madal, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, R.  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miranda, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A.  
Rodrigues, V.; Rosa, A.J.; de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Savasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tsuchiko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2445

Query Match 68.6%; Score 35; DB 2; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 92;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 GRIIPVGR 10  
|:|:|:|:  
Db 1047 GLRPVGR 1053

RESULT 9

A65057  
hypothetical protein b2757 - *Escherichia coli* (strain K-12)  
C:Species: *Escherichia coli*  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
C:Accession: A65057  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of *Escherichia coli* K-12.  
A:Reference number: A64720; M0ID:97426617  
A:Accession: A65057  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-226 <BLAT>  
A:Cross-references: GB:AE000359; GB:U00096; NID:g1789110; PIDN:AACT5799.1; PID:g17891  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: *Escherichia coli* hypothetical protein b2757

Query Match 66.7%; Score 34; DB 2; Length 226;  
Best Local Similarity 75.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GRIIPVGR 11  
|:|:|:|:  
Db 25 GRIIPVGR 32

RESULT 10  
S51485  
ribosomal protein S1 - *Synechococcus* sp. (PCC 6301)  
C:Species: *Synechococcus* sp.  
A:Variety: PCC 6301  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000  
C:Accession: S51485  
R:Sugita, M.; Sugita, C.; Sugita, M.  
Mol. Gen. Genet. 246, 142-147, 1995  
A:Title: Structure and expression of the gene encoding ribosomal protein S1 from the  
n CSI.  
A:Reference number: S51483; M0ID:95166170  
A:Accession: S51485  
A:Molecule type: DNA  
A:Residues: 1-307 <SUG>  
A:Cross-references: EMBL:D8752; NID:g560122; PIDN:BA05946.1; PID:g666973  
A:Experimental source: PCC 6301  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: *Synechocystis* ribosomal protein S1  
C:Keywords: protein biosynthesis; ribosome

Query Match 66.7%; Score 34; DB 2; Length 307;  
Best Local Similarity 66.7%; Pred. No. 41;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 GRIIPVGR 11  
|:|:|:|:  
Db 205 GRIIPVGR 213

RESULT 11  
S77236  
ribosomal protein S1 - *Synechocystis* sp. (strain PCC 6803)  
N:Alternate names: protein slr1356  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S77236  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O.; Kanehisa, S.; Shimizu, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys*  
s.

A:Reference number: S74322; MUID:97061201  
A:Accession: S77236  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-328 <KAN>  
A:Cross-references: EMBL:DP0907; GB:AB001339; NID:g1652618; PIDN:BAI1570.1; PID:g165265  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: Synecocystis ribosomal protein S1  
C:Keywords: protein biosynthesis; ribosome

Query Match 66.7%; Score 34; DB 2; Length 328;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||:|:|  
DB 204 RGIRPVGAR 212

RESULT 12  
AH1823  
30S ribosomal protein S1 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH1823  
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kiritz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi, N.; Nakazaki, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH1823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-343 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BAJ77660.1; PID:g17135114; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: Synecocystis ribosomal protein S1

Query Match 66.7%; Score 34; DB 2; Length 343;  
Best Local Similarity 66.7%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||:|:|  
DB 205 RGIRPVGAR 213

RESULT 13  
P97646  
probable aliphatic sulfonates transport permease protein ssuc AGR\_C.4335 [imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: P97646  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Gurollo, B.; Goldman, A.; Liu, F.; Mollm, C.; Allinger, M.; Doughy, D.; Scott, C.; Lappas, C.; Matkeitz, B.; Science 294, 2223-2320, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; PMID:11743194  
A:Accession: P97646  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK8127.1; PID:g15157563; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C.4335  
A:Map position: circular chromosome

Query Match 64.7%; Score 33; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||:|:|  
DB 123 RGLPIGRNF 131

RESULT 14  
AD2870  
ABC transporter, membrane spanning protein Atu2390 [imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AD2870  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woeherage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl ; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AD2870  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE006888; PIDN:AAU43378.1; PID:g17740875; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu2390  
A:Map position: circular chromosome

Query Match 64.7%; Score 33; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 3 RGIRPVGRF 11  
|||:|:|  
DB 123 RGLPIGRNF 131

RESULT 15  
B95351  
virB type IV secretion protein [imported] - Sinorhizobium meliloti (strain 1021) mag  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: B95351  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B ; Kaiman, S.; Keating, D.H.; Palm, M.C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A:Reference number: A95262; MUID:21396509; PMID:11481452  
A:Accession: B95351  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AAK5372.1; PID:g14523833; GSPDB:GN00165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Flhau, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, D.; Chailin, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kaiman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lejau Hebaull, P.; Vandembol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: virB6

A:Genome: plasmid

Query Match 64.7%; Score 33; DB 2; Length 313;  
Best Local Similarity 75.0%; Pred. No. 66;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 3 RGIRPVGR 10  
||:||||  
Db 297 RGVRAVGR 304

Search completed: September 13, 2002, 09:24:02  
Job time: 777 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:47 ; Search time 80.21 seconds  
(without alignments)  
5.793 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_10\_21

Perfect score: 51  
Sequence: 1 XXRGIRPVGRFX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	DB ID	Description
1	48	94.1	83 1	PRRP_RAT
2	48	94.1	87 1	PRRP_HUMAN
3	48	94.1	98 1	PRRP_BOVIN
4	35	68.6	391 1	THIL_ZOORA
5	35	68.6	922 1	GYRA_AERSEA
6	34	66.7	224 1	YGC1_ECOLI
7	34	66.7	306 1	RSL_SYNP6
8	34	66.7	328 1	RS1A_SYNY3
9	34	66.7	801 1	TFR2_HUMAN
10	33	64.7	477 1	PPOX_HUMAN
11	33	64.7	477 1	PPOX_MOUSE
12	33	64.7	485 1	SYE_BACHD
13	32	62.7	150 1	ULAZ_HCMVA
14	32	62.7	249 1	Y361_SYNY3
15	32	62.7	330 1	AR72_HUMAN
16	32	62.7	331 1	AR73_HUMAN
17	32	62.7	676 1	EXL1_HUMAN
18	31	60.8	130 1	RL3_PIG
19	31	60.8	149 1	ENRN_BPT7
20	31	60.8	207 1	YD05_SCHPO
21	31	60.8	220 1	PYRF_HABIN
22	31	60.8	233 1	PYRF_PASPU
23	31	60.8	250 1	LINC_PSEPA
24	31	60.8	364 1	MK11_MOUSE
25	31	60.8	367 1	GCST_MYCTU
26	31	60.8	372 1	MK11_HUMAN
27	31	60.8	413 1	CSD_THEMA
28	31	60.8	450 1	PYRC_MYCTU
29	31	60.8	433 1	ZU01_YEAST
30	31	60.8	463 1	K10_DROME
31	31	60.8	625 1	DHGL_DROME
32	31	60.8	625 1	DHGL_DROME
33	31	60.8	631 1	DNAK_THERO

34	31	60.8	719 1	NRP1_YEAST	P32770 saccharomyc
35	31	60.8	908 1	ACON_BACSU	P09339 bacillus su
36	31	60.8	1056 1	POL_BIV06	P19560 bovine immu
37	31	60.8	1056 1	POL_BIV27	P19561 bovine immu
38	31	60.8	1416 1	BIM_MOUSE	O68700 mus musculu
39	31	60.8	1561 1	SPAP_STRMU	P23504 streptococc
40	31	60.8	1565 1	PAC_STRMU	P11657 streptococc
41	30	58.8	116 1	RL34_HUMAN	P49207 homo sapien
42	30	58.8	179 1	RK6_GUTH	O46908 guillardi
43	30	58.8	292 1	CNTB_HUMAN	O9uff9 homo sapien
44	30	58.8	368 1	YG3_YEAST	P53295 saccharomyc
45	30	58.8	376 1	MPK5_ARATH	O39025 arabidopsis

## ALIGNMENTS

```

RESULT 1
PRRP_RAT STANDARD; PRT; 83 AA.
AC P81278;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Prolactin-releasing peptide precursor (PrpP) (Prolactin-releasing
DE hormone) [Contains: Prolactin-releasing peptide PrpP31; Prolactin-
DE releasing peptide PrpP20].
GN PRP.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; Pubmed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RP TISSUE SPECIFICITY.
RX Pubmed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PrpP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -1- FUNCTION: Stimulates prolactin (PrL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PrL.
CC -1- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC medulla oblongata and hypothalamus.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: AB015418; BAA29026.1;
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1
FT PEPTIDE 21
FT PEPTIDE 22 52
FT PEPTIDE 33 52
FT MOD_RES 52 52
FT SEQUENCE 83 AA; 9215 MW; DDC75A264DEBAF29 CRC64;

```

Query Match 94.1%; Score 48; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.004;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
 |||||  
 DB 44 RGIRPYGRF 52

RESULT 2  
 ID PRRP\_HUMAN STANDARD: PRT; 87 AA.  
 AC P81277;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].  
 GN PRH.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Brain;  
 RX MEDLINE=9826781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 RN [2]  
 RP TISSUE SPECIFICITY.  
 RX PubMed=10498338;  
 RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its receptor.";  
 RL Regul. Pept. 83:1-10(1999).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; AB015419; BAA29027.1; -  
 KW Hormone; Amidation; Signal.  
 FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE 34 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

Query Match 94.1%; Score 48; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
 |||||  
 DB 45 RGIRPYGRF 53

RESULT 3  
 ID PRRP\_BOVIN STANDARD: PRT; 98 AA.  
 AC P81264;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PrRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].  
 GN PRH.  
 OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.  
 OX NCBI\_TaxID=9913;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RC TISSUE=Brain;  
 RX MEDLINE=9826781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A prolactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 DR EMBL; AB015417; BAA29025.1; -  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA508 CRC64;

Query Match 94.1%; Score 48; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
 |||||  
 DB 45 RGIRPYGRF 53

RESULT 4  
 ID THIL\_ZOORA STANDARD: PRT; 391 AA.  
 AC P07057;  
 DT 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Acetyl-CoA acetyltransferase (EC 2.3.1.9) (Acetoacetyl-CoA thiolase).  
 GN PHBA.  
 OS Zoogloea ramifera.  
 OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group; Zoogloea.  
 OX NCBI\_TaxID=350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19623 / I-16-M;

```

RX MEDLINE-87083504; PubMed-2878929;
RA Peoples O.P., Masamune S., Walsh C.T., Sinskey A.J.;
RT Biosynthetic thiolase from Zoogloea ramigera. III. Isolation and
RL J. Biol. Chem. 262:97-102(1987).
RN [2]
RP REVISION TO 130
RX MEDLINE-89359356; PubMed-2670935;
RA Peoples O.P., Sinskey A.J.;
RT "Poly-beta-hydroxybutyrate biosynthesis in Alcaligenes eutrophus H16.
RT Characterization of the genes encoding beta-ketothiolase and
RL acetacetyl-CoA reductase."
RL J. Biol. Chem. 264:15293-15297(1989).
RN [3]
RP MUTAGENESIS OF CYS-377.
RX MEDLINE-91217075; PubMed-1673680;
RA Palmer M.A.J., Differding E., Gamboni R., Williams S.F., Peoples O.P.,
RA Walsh C.T., Sinskey S.J., Masamune S.;
RT "Biosynthetic thiolase from Zoogloea ramigera. Evidence for a
RT mechanism involving Cys-378 as the active site base."
RL J. Biol. Chem. 266:8369-8375(1991).
CC -1- CATALYTIC ACTIVITY: 2 acetyl-CoA + CoA + acetoacetyl-CoA.
CC -1- PATHWAY: FIRST STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02631; AAA27706.1; ALT_SEQ.
DR PIR: A26121; XKGZAC.
DR PIR: A27754; A27754.
DR HSSP: P27796; IAFY.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase_1.
DR Pfam: PF02803; thiolase_1.
DR PROSITE: PS00098; THIOLEASE_1.
DR PROSITE: PS00099; THIOLEASE_3.
DR PROSITE: PS00737; THIOLEASE_2.
DR TRANSFERASE: ACYLtransferase; PHB biosynthesis.
FT INIT_MET 0 0
FT ACT_SITE 88 88 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 377 377 BASE.
FT MUTAGEN 377 377 C->G: LOSS OF ACTIVITY.
SQ SEQUENCE 391 AA; 40342 MW; 6D2351A1BC0E4BDD CRC64;

```

```

Query Match 68.6%; Score 35; DB 1; Length 391;
Best Local Similarity 75.0%; Pred. No. 9.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 RGIRPVR 10
Db 266 RGIRPVR 273

```

```

RESULT 5
GYRA_AERSA STANDARD; PRT; 922 AA.
AC P48369;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA gyrase subunit A (EC 5.99.1.3).
GN GYRA.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;

```

```

OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2148/89;
RA Opegaard H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 33-179 FROM N.A.
RC STRAIN=ATCC 14174;
RX MEDLINE-95142596; PubMed-7840589;
RA Opegaard H., Sorum H.;
RT "Gyrase mutations in quinolone-resistant isolates of the fish pathogen
RT Aeromonas salmonicida."
RL Antimicrob. Agents Chemother. 38:2460-2464(1994).
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING. THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: I47978; AAB41037.1; -.
DR EMBL: I42453; AAA87239.1; -.
DR HSSP: P09097; IAB4.
DR InterPro: IPR002205; DNA_topoisomIV.
DR Pfam: PF00521; DNA_topoisomIV.
DR SMART: SM00434; TOP4c_1.
DR Isomerase: Topoisomerase; DNA-binding.
FT ACT_SITE 122 122 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 922 AA; 101333 MW; 8894965DC4217077 CRC64;

```

```

Query Match 68.6%; Score 35; DB 1; Length 922;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 RGIRPVR 10
Db 758 KGVPRMGR 765

```

```

RESULT 6
YGC1_ECOLI STANDARD; PRT; 224 AA.
ID YGC1_ECOLI
AC Q46898;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ygc1 precursor.
GN YGC1 OR B2757.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

RA Mau B., Shao Y.,  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U29579; AAA69267.1; ALT\_INIT.  
DR EMBL: AE000359; AAC75799.1; ALT\_INIT.  
DR Ecocore; EG13116; YGCI.  
KW Hypothetical protein; Signal; Complete proteome.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 224 HYPOTHETICAL PROTEIN YGCI.  
SQ SEQUENCE 224 AA: 25209 MW: 1C42CC009B317D68 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 224;  
Best Local Similarity 75.0%; Pred. No. 8.5;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 4 GIRPYGRF 11  
DB 23 GTRPTGRF 30

RESULT 7  
RS1\_SYNP6 STANDARD; PRT; 306 AA.  
AC P46228;  
DT 01-NOV-1995 (Rel. 32; Created)  
DT 01-FEB-1996 (Rel. 33; Last sequence update)  
DT 16-OCT-2001 (Rel. 40; Last annotation update)  
DE 30S ribosomal protein S1.  
GN RPSA OR RPS1.  
OS Synecchococcus sp. (strain PCC 6301) (Anacystis nidulans).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecchococcus.  
OX NCBI\_TaxID=1139;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-13.  
RA MEDLINE-9516170; PubMed-7862084;  
RA Sugita M., Sugita C., Sugita M.,  
RT "Structure and expression of the gene encoding ribosomal protein S1  
RT from the cyanobacterium Synecchococcus sp. strain PCC 6301: striking  
RT sequence similarity to the chloroplast ribosomal protein S1.";  
RL Mol. Gen. Genet. 246:142-147(1995).  
-----  
CC -1- FUNCTION: BINDS MRNA.  
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 3 S1 MOTIF DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: D28752; BAA05946.1; -.  
DR HSP; P05055; ISRO.  
DR InterPro; IPR003029; S1.  
DR Pfam; PF00575; S1; 3.  
DR SMART; SM00316; S1; 3.  
DR PROSITE; PS50126; S1; 3.  
KW Ribosomal protein; Repeat; RNA-binding.  
FT INIT\_MET 0  
FT DOMAIN 31 100 S1 MOTIF 1.  
FT DOMAIN 118 182 S1 MOTIF 2.  
FT DOMAIN 196 264 S1 MOTIF 3.

SQ SEQUENCE 306 AA: 34492 MW: 82D4A2EBE69B7CE CRC64;

Query Match 66.7%; Score 34; DB 1; Length 306;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
DB 204 RGIRPYGAF 212

RESULT 8  
RS1A\_SYNY3 STANDARD; PRT; 328 AA.  
ID RS1A\_SYNY3  
AC P73530;  
DT 01-NOV-1997 (Rel. 35; Created)  
DT 01-NOV-1997 (Rel. 35; Last sequence update)  
DT 16-OCT-2001 (Rel. 40; Last annotation update)  
DE 30S ribosomal protein S1 homolog A.  
GN RPS1A OR SLR1356.  
OS Synecchocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecchocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97061201; PubMed-8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirose M., Sugita M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naro K.,  
RA Yamada M., Yasuda M., Tabata S.,  
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecchocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-116(1996).  
-----  
CC -1- FUNCTION: BINDS MRNA.  
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 3 S1 MOTIF DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: D90907; BAA17570.1; -.  
DR HSP; P05055; ISRO.  
DR InterPro; IPR003029; S1.  
DR Pfam; PF00575; S1; 3.  
DR SMART; SM00316; S1; 3.  
DR PROSITE; PS50126; S1; 3.  
KW Ribosomal protein; Repeat; RNA-binding; Complete proteome.  
FT DOMAIN 31 100 S1 MOTIF 1.  
FT DOMAIN 118 182 S1 MOTIF 2.  
FT DOMAIN 196 264 S1 MOTIF 3.  
SQ SEQUENCE 328 AA: 36570 MW: DC3FF9B1E5A40619 CRC64;

Query Match 66.7%; Score 34; DB 1; Length 328;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 3 RGIRPYGRF 11  
DB 204 RGIRPYGAF 212

RESULT 9  
TFR2\_HUMAN STANDARD; PRT; 801 AA.  
ID TFR2\_HUMAN

AC Q9UP52: 075422; Q9HA99; Q9N67;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DE 01-MAR-2002 (Rel. 41, Last annotation update)  
 GN Transferrin receptor protein 2 (TFR2).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A. (ISOFORM ALPHA AND BETA).  
 RC TISSUE-Erythroid leukemia, and myeloid leukemia cells;  
 RX MEDLINE-99340005; PubMed-10409623;  
 RA Kawabata H., Yang R., Hirama T., Vuong P.T., Kawano S., Gombart A.F.,  
 RA Koeffler H.P.;  
 RT "Molecular cloning of transferrin receptor 2: a new member of the  
 RT transferrin receptor-like family".  
 RL J. Biol. Chem. 274:20826-20832(1999).  
 RN [2]  
 RP SEQUENCE FROM N.A. (ISOFORM GAMMA).  
 RX PubMed-9799793;  
 RA Gloeckner G., Scherer S., Schatteroy R., Borlight A.P., Weber J.,  
 RA Tsai L.-C., Rosenthal A.;  
 RT "Large scale analysis of two regions in human chromosome 7q22:  
 RT annotation of 650 kb of genomic sequence around the EPO and CUTL1 loci  
 RT reveals 17 genes.";  
 RL Genome Res. 8:1060-1073(1998).  
 RN [3]  
 RP SEQUENCE OF 1-158 AND 370-801 FROM N.A.  
 RC TISSUE-Carcinoma, and Embryo;  
 RA Itoigai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
 RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
 RA Wagaitsuma M., Hosokiri T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
 RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takichuchi S.,  
 RA Watanabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
 RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,  
 RA Niimiya K., Iwayanagi T.;  
 RT "NEO human cDNA sequencing project.";  
 RN Submitted (Aug-2000) to the EMBL/Genbank/DBJ databases.  
 RP DISEASE:  
 RX PubMed-10802645;  
 RA Camaschella C., Roetto A., Call A., De Gobbi M., Garozzo G.,  
 RA Carella M., Majorano N., Totaro A., Gasparini P.;  
 RT "The gene TFR2 is mutated in a new type of haemochromatosis mapping to  
 RT 7q22.";  
 RL Nat. Genet. 25:14-15(2000).  
 RN [5]  
 RP VARIANT HFE3 LYS-172.  
 RX PubMed-11313241;  
 RA Roetto A., Totaro A., Piperno A., Piga A., Longo F., Garozzo G.,  
 RA Call A., De Gobbi M., Gasparini P., Camaschella C.;  
 RT "New mutations inactivating transferrin 2 in hemochromatosis type  
 RT 3.";  
 RL Blood 97:2555-2560(2001).  
 CC -1- FUNCTION: Mediates cellular uptake of transferrin-bound iron in a  
 CC non-iron dependent manner. May be involved in iron metabolism,  
 CC hepatocyte function and erythrocyte differentiation.  
 CC -1- SUBUNIT: Homodimer.  
 CC -1- SUBCELLULAR LOCATION: Type II membrane protein. The beta isoform  
 CC lacks the transmembrane domain and is probably intracellular.  
 CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; alpha (shown here), beta and  
 CC gamma; are produced by alternative splicing.  
 CC -1- TISSUE SPECIFICITY: Predominantly expressed in liver. While the  
 CC alpha form is also expressed in spleen, lung, muscle, prostate and  
 CC peripheral blood mononuclear cells, the beta form is expressed in  
 CC all tissues tested, albeit weakly.  
 CC -1- DISEASE: Defects in TFR2 are a cause of hereditary hemochromatosis  
 CC type III (HFE3). HFE3 is a disorder of iron homeostasis resulting  
 CC in iron overload and has a phenotype indistinguishable from that  
 CC of hereditary hemochromatosis (HH). HH is characterized by  
 CC abnormal intestinal iron absorption and progressive increase of

CC total body iron, which results in midlife in clinical  
 CC complications including cirrhosis, cardiopathy, diabetes,  
 CC endocrine dysfunctions, arthropathy, and susceptibility to liver  
 CC cancer. Since the disease complications can be effectively  
 CC prevented by regular phlebotomies, early diagnosis is most  
 CC important to provide a normal life expectancy to the affected  
 CC subjects.  
 CC -1- MISCELLANEOUS: The variant lys-172 found in hereditary  
 CC hemochromatosis type III affects the putative initiation codon of  
 CC the beta isoform thus preventing its translation.  
 CC -1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2B8.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL: AF067864; AAD45561.1; -;  
 DR EMBL: AF053356; AAC78796.1; -;  
 DR EMBL: AK022002; BAB13951.1; -;  
 DR EMBL: AK000421; BAB91153.1; ALT\_INIT.  
 DR MEROPS; M28.973; -;  
 DR MIM; 604720; -;  
 DR MIM; 604250; -;  
 DR InterPro: IPR003137; PA.  
 DR Pfam: PF02225; PA. 1  
 KW Transmembrane; Glycoprotein; Receptor; Signal-anchor;  
 KW Alternative splicing; Disease mutation.  
 FT DOMAIN 1 83  
 FT TRANSMEM 84 104  
 FT FT  
 FT DOMAIN 105 801  
 FT SITE 23 26  
 FT DISULFID 108 108  
 FT DISULFID 111 111  
 FT CARBOHYD 240 240  
 FT CARBOHYD 339 339  
 FT CARBOHYD 540 540  
 FT CARBOHYD 754 754  
 FT CARBOHYD 1 171  
 FT VARSPLIC 343 369  
 FT VARIANT 172 172  
 FT FT  
 FT CONFLICT 712 712 R -> RIPSNOV (IN REF. 2).  
 FT SEQUENCE 801 AA; 88755 MW; D3D3082BA835413A CRC64;  
 SQ  
 QY 3 RGIRPYGR 10  
 DB 262 RGVDPVGR 269  
 DB 11:1111  
 Query Match 66.7%; Score 34; DB 1; Length 801;  
 Best local Similarity 75.0%; Pred. No. 31;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 RESULT 10  
 ID PPOX\_HUMAN STANDARD; PRT; 477 AA.  
 AC P50336;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Protoporphyrinogen oxidase (EC 1.3.3.4) (PPO).  
 GN PPOX.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]

RP SEQUENCE FROM N.A.  
RC TISSUE-Placenta;  
RX MEDLINE-95229621; PubMed-7713909;  
RA Nishimura K., Taketani S., Inokuchi H.;  
RT "Cloning of a human cDNA for protoporphyrinogen oxidase by  
complementation in vivo of a hemg mutant of Escherichia coli.";  
RL J. Biol. Chem. 270:8076-8080(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Placenta;  
RX MEDLINE-96367087; PubMed-8771201;  
RA Dailey T.A., Dailey H.A.;  
RT "Human protoporphyrinogen oxidase: expression, purification, and  
characterization of the cloned enzyme.";  
RL Protein Sci. 5:98-105(1996).  
RN [3]  
RP VARIANT VP ARG-232, AND VARIANT HIS-304.  
RX MEDLINE-97005368; PubMed-8852667;  
RA Deybach J.-C., Puy H., Rodreau A.-M., Lamort J., da Silva V.,  
Grandchamp B., Nordmann Y.;  
RT "Mutations in the protoporphyrinogen oxidase gene in patients with  
variegate porphyria.";  
RL Hum. Mol. Genet. 5:407-410(1996).  
RN [4]  
RP VARIANT VP TRP-59, AND VARIANT CYS-168.  
RX MEDLINE-96241580; PubMed-8673113;  
RA Meisner P.N., Dailey T.A., Hift R.J., Ziman M., Corrigan A.V.,  
Roberts A.G., Meisner D.M., Kirsch R.E., Dailey H.A.;  
RT "A R59W mutation in human protoporphyrinogen oxidase results in  
decreased enzyme activity and is prevalent in South Africans with  
variegate porphyria.";  
RL Nat. Genet. 13:95-97(1996).  
RN [5]  
RP VARIANT VP CYS-152.  
RX MEDLINE-98434271; PubMed-9763307;  
RA Frank J., Poh-Fitzpatrick M.B., King L.E. Jr., Cristiano A.M.;  
RT "The genetic basis of 'Scarsdale Gourmet Diet' variegate porphyria: a  
missense mutation in the protoporphyrinogen oxidase gene.";  
RL Arch. Dermatol. Res. 290:441-445(1998).  
CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRIN  
IX TO FORM PROTOPORPHYRIN IX.  
CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) -> protoporphyrin-  
IX + H(2)O(2).  
CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.  
CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
CC -1- SUBUNIT: HOMODIMER.  
CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE  
WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).  
CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,  
LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.  
CC -1- DISEASE: DEFECTS IN PPOX ARE THE CAUSE OF PORPHYRIA VARIEGATA  
(VP), A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR  
HYPERTRICHOSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE  
INTERMITTENT PORPHYRIA.  
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRIN OXIDASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D38537; BAA07538.1; -  
DR EMBL: U26446; AAA67690.1; -  
DR MIM: 600923; -  
DR MIM: 176200; -  
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;  
KW Flavoprotein; FAD; Mitochondrion; Disease mutation; Polymorphism.  
FT NP\_BIND 9 14 FAD (POTENTIAL).  
FT VARIANT 59 59 R -> W (IN VP).  
FT /FTid=VAR\_003686.

FT VARIANT 152 152 R -> C (IN VP).  
FT /FTid=VAR\_003687.  
FT VARIANT 168 168 R -> C.  
FT /FTid=VAR\_003688.  
FT VARIANT 232 232 G -> R (IN VP).  
FT /FTid=VAR\_003689.  
FT VARIANT 304 304 R -> H.  
FT /FTid=VAR\_003690.  
SQ SEQUENCE 477 AA; 50765 MW; 2444DEAC2E6C33EE CRC64;  
  
Query Match 64.7%; Score 33; DB 1; Length 477;  
Best Local Similarity 85.7%; Pred. NO. 29;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 3 RGIRPVG 9  
DB 59 RGIRPVG 65  
  
RESULT 11  
PPOX\_MOUSE  
ID PPOX\_MOUSE STANDARD; PRT: 477 AA.  
AC P51175; P97344;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 15-DEC-1998 (Rel. 37, Last annotation update)  
DE Protoporphyrinogen oxidase (EC 1.3.3.4) (PPO).  
GN PPOX.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
OX NCBI\_Taxid-10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-96132666; PubMed-8554330;  
RA Dailey T.A., Dailey H.A., Meisner P., Prasad A.R.;  
RT "Cloning, sequence, and expression of mouse protoporphyrinogen  
oxidase.";  
RL Arch. Biochem. Biophys. 324:379-384(1995).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE-Erythrocyte;  
RX MEDLINE-95331315; PubMed-7607249;  
RA Taketani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,  
Nishimura K., Inokuchi H.;  
RT "Induction of terminal enzymes for heme biosynthesis during  
differentiation of mouse erythrocyte cells.";  
RL Eur. J. Biochem. 230:760-765(1995).  
RN [3]  
RP SUBCELLULAR LOCATION.  
RC TISSUE-Liver;  
RX MEDLINE-88153682; PubMed-3346226;  
RA Ferreira G.C., Andrew T.L., Kerr S.W., Dailey H.A.;  
RT "Organization of the terminal two enzymes of the heme biosynthetic  
pathway. Orientation of protoporphyrinogen oxidase and evidence for a  
membrane complex.";  
RL J. Biol. Chem. 263:3835-3839(1988).  
CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRIN  
IX TO FORM PROTOPORPHYRIN IX.  
CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) -> protoporphyrin-  
IX + H(2)O(2).  
CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).  
CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE  
WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.  
CC -1- INDUCTION: DURING ERYTHROID DIFFERENTIATION.  
CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRIN OXIDASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U25114; AAA6003.1; -  
DR EMBL: D45185; BAA08126.1; -  
DR MGD: MGI:104968; Ppx.  
KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;  
KW Flavoprotein; FAD; Mitochondrion.  
FT NP\_BIND 9 14 PAD (POTENTIAL).  
FT CONFLICT 64 64 A -> T (IN REF. 2).  
FT CONFLICT 66 66 A -> P (IN REF. 2).  
FT CONFLICT 108 108 L -> S (IN REF. 2).  
FT CONFLICT 427 427 W -> C (IN REF. 2).  
SQ SEQUENCE 477 AA; 50870 MW; 8CFB48120728DE6F CRC64;

Query Match 64.7%; Score 33; DB 1; Length 477;  
Best Local Similarity 85.7%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RGIRPVG 9  
Db 59 RGIRPAG 65

RESULT 12  
SYE\_BACHD  
ID SYE\_BACHD STANDARD; PRT; 485 AA.  
AC 09KGR6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Glutamy-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)  
DE (GLURS)  
GN GLT3 OR BH0109.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
Fuji F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
halodurans and genomic sequence comparison with Bacillus subtilis."  
RL Nucleic Acids Res. 28:4317-4331(2000).  
CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) -> AMP +  
diphosphate + L-glutamy-tRNA(Glu).  
CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOCYL-TRNA SYNTHETASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AF001507; BAB03828.1; -  
DR InterPro: IPR000924; tRNA-synt\_1c.  
DR InterPro: IPR001412; tRNA-synt\_1c.  
DR Pfam: PF00749; tRNA-synt\_1c; 1.  
DR PRINTS: PR00987; TRNASYNTHGU.  
DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I; 1.  
KW Aminocyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
KW Complete proteome.

FT SITE 11 21 "HIGH" REGION.  
FT SITE 252 256 "RMSK" REGION.  
FT BINDING 255 255 ATP (BY SIMILARITY).  
SQ SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match 64.7%; Score 33; DB 1; Length 485;  
Best Local Similarity 66.7%; Pred. No. 30;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 RGIRPVGRF 11  
Db 149 RGIRPVGRF 157

RESULT 13  
UL47\_HCMVA  
ID UL47\_HCMVA STANDARD; PRT; 150 AA.  
AC P16828;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Hypothetical protein UL107.  
GN UL107.  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
OX NCBI\_TaxID=10360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=90269039; PubMed=2161319;  
RA Chee M.S., Bankier A.T., Beck S., Bohnl R., Brown C.M., Cerny R.,  
Horsnell T., Hutchison C.A. III, Kouzarides T., Martignetti J.A.,  
Preddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
RT "Analysis of the protein-coding content of the sequence of human  
cytomegalovirus strain AD169."  
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X17403; CAA35343.1; -  
DR PIR: S09872; S09872.  
KW Hypothetical protein.  
FT CARBOHYD 144 144 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 150 AA; 17373 MW; 042707546C7EB878 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 150;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 3 RGIRPVGRF 11  
Db 57 RGIRPVGRF 65

RESULT 14  
Y361\_SYNY3  
ID Y361\_SYNY3 STANDARD; PRT; 249 AA.  
AC Q35378;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 28.2 kDa protein slr0361.  
GN slr0361.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

```

OX NCBI_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96127529; PubMed=8559279;
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugita M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; D63999; BAA10082.1; -
DR InterPro: IPR000613; PseudoU_synth.
DR InterPro: IPR000748; Psi_RSU.
DR InterPro: IPR002942; S4.
DR Pfam; PF00849; PseudoU_synth_2; 1.
DR Pfam; PF01479; S4; 1.
DR SMART; SM00363; S4; 1.
DR PROSITE; PS01149; Psi_RSU; 1.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 249 AA; 28228 MW; 315F208030180326 CRC64;

Query Match 62.7%; Score 32; DB 1; Length 249;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 3 GIRPVGR 10
DB 100 KGLRPVGR 107

RESULT 15
AR72_HUMAN STANDARD; PRT; 330 AA.
AC 043488; 075749;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aflatoxin B1 aldehyde reductase 1 (Ec 1.-.-.-) (AFB1-AR 1)
DE (Aldoketoreductase 7).
GN AKR7A2 OR AFAR OR AKR7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Liver;
RA MEDLINE=98244807; PubMed=9576847;
RA Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;
RT "Molecular cloning, expression and catalytic activity of a human AKR7
RT member of the aldo-keto reductase superfamily: evidence that the
RT major 2-carboxybenzaldehyde reductase from human liver is a homologue
RT of rat aflatoxin B1-aldehyde reductase.";
RL Biochem. J. 332:21-34(1998).
RN [2]
RP SEQUENCE FROM N.A.
RP TISSUE=Brain;
RA MEDLINE=99040634; PubMed=9823300;
RA Pramli C., Savelyeva L., Perri P., Schwab M.;
RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-
RT 1p36.1 in a region frequently altered in human tumor cells.";
RL Cancer Res. 58:5014-5018(1998).

```

```

RN [3]
RP SEQUENCE FROM N.A.
RA Hall R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIACETOL. COULD BE
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC
CC EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE
CC REDUCTASE.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF026947; AAC52104.1; -
DR EMBL; Y16675; CAA76347.1; -
DR EMBL; AL035413; CAB72321.1; -
DR MIM; 603418; -
DR InterPro: IPR001395; Aldo_ket_red.
DR Pfam; PF00248; aldo_ket_red; 1.
DR Oxidoreductase.
KW ACT_SITE 112 HYDROGEN-BOND DONOR (PROBABLE).
FT ACCT_SITE 112 A->T (IN REF. 1).
FT CONFLICT 113
FT SEQUENCE 330 AA; 36618 MW; 3BBF7ED0CA4D54 CRC64;

```

```

Query Match 62.7%; Score 32; DB 1; Length 330;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GIRPVGRF 11
DB 215 GKRPVGRF 222

```

Search completed: September 13, 2002, 09:30:48  
Job time: 1138 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:23 ; Search time 311.85 Seconds  
(without alignments)  
6.657 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_10\_21  
Perfect score: 51  
Sequence: 1 XXRGIRPVGRFX 12

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp\_archaea:\*  
2: sp\_bacteria:\*  
3: sp\_fungi:\*  
4: sp\_human:\*  
5: sp\_invertebrate:\*  
6: sp\_mammal:\*  
7: sp\_mhc:\*  
8: sp\_organelle:\*  
9: sp\_phage:\*  
10: sp\_plant:\*  
11: sp\_rodent:\*  
12: sp\_virus:\*  
13: sp\_vertebrate:\*  
14: sp\_unclassified:\*  
15: sp\_virus:\*  
16: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	90.2	117	13	09W624
2	43	84.3	420	2	0931z7
3	36	70.6	426	17	097950
4	36	70.6	674	16	050431
5	35	68.6	272	3	042924
6	35	68.6	371	16	0981v3
7	35	68.6	504	16	053895
8	35	68.6	519	16	092567
9	35	68.6	894	16	09KSJ8
10	35	68.6	1000	5	09VS40
11	35	68.6	1056	16	09PAQ0
12	35	68.6	1172	5	09VU79
13	34	66.7	329	2	09RJ77
14	34	66.7	529	3	09C1M8
15	34	66.7	537	4	075230
16	34	66.7	780	4	075422

17	34	66.7	801	4	09UP52	09UP52 homo sapien
18	33	64.7	139	5	09N6S1	09N6S1 leishmania
19	33	64.7	153	3	09T133	09T133 bacterioph
20	33	64.7	205	4	096FC9	096FC9 homo sapien
21	33	64.7	253	2	P72459	P72459 streptomyce
22	33	64.7	313	16	0921Z8	0921Z8 rhizobium m
23	33	64.7	319	2	088067	088067 streptomyce
24	33	64.7	335	16	P96221	P96221 mycobacteri
25	33	64.7	353	16	031194	031194 mycobacteri
26	33	64.7	366	16	066950	066950 aquifex neo
27	33	64.7	445	2	050531	050531 streptomyce
28	33	64.7	477	11	099W34	099W34 mus musculu
29	33	64.7	511	3	059714	059714 schizosacch
30	33	64.7	589	16	09A718	09A718 caulobacter
31	33	64.7	883	10	09LE66	09LE66 lycoperisoc
32	33	64.7	901	16	099YPS	099YPS streptococc
33	33	64.7	1940	16	09NRC8	09NRC8 deinococcus
34	33	64.7	2104	5	021281	021281 caenorhabdl
35	33	64.7	2104	5	0964N4	0964N4 caenorhabdl
36	33	62.7	97	8	09B5K5	09B5K5 balanus gla
37	32	62.7	154	2	054139	054139 streptomyce
38	32	62.7	183	16	092KK8	092KK8 rhizobium m
39	32	62.7	183	17	0979Z6	0979Z6 thermoplasm
40	32	62.7	286	8	023888	023888 dictyostell
41	32	62.7	278	16	092OR1	092OR1 rhizobium m
42	32	62.7	285	2	09X4W7	09X4W7 pseudomonas
43	32	62.7	292	16	092VM1	092VM1 rhizobium m
44	32	62.7	299	16	P95060	P95060 mycobacteri
45	32	62.7	309	2	086546	086546 streptomyce

## ALIGNMENTS

RESULT	ID	Query Match	Length	DB ID	Description
1	09W624	90.2	117	13	09W624
2	09W624	84.3	420	2	0931z7
3	09W624	70.6	426	17	097950
4	09W624	70.6	674	16	050431
5	09W624	68.6	272	3	042924
6	09W624	68.6	371	16	0981v3
7	09W624	68.6	504	16	053895
8	09W624	68.6	519	16	092567
9	09W624	68.6	894	16	09KSJ8
10	09W624	68.6	1000	5	09VS40
11	09W624	68.6	1056	16	09PAQ0
12	09W624	68.6	1172	5	09VU79
13	09W624	66.7	329	2	09RJ77
14	09W624	66.7	529	3	09C1M8
15	09W624	66.7	537	4	075230
16	09W624	66.7	780	4	075422

Query Match	Score	DB	Length	Indels	Gaps
Best Local Similarity	77.8	Pred. No. 0.11			
Matches 7; Conservative	2	Mismatches 0			
QY	3	RGIRPVGRF 11			
Db	67	RGVPRIGRF 75			
RESULT	2				
ID	0931z7	PRELIMINARY;	PRT;	420 AA.	
AC	0931z7				
DT	01-DEC-2001	(TREMBLrel. 19, Created)			
DT	01-DEC-2001	(TREMBLrel. 19, last sequence update)			
DT	01-DEC-2001	(TREMBLrel. 19, last annotation update)			

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
 OS AUR2B.  
 GN Streptomyces aureofaciens.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1894;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CCM3239;  
 RA Kornanev J., Bistakova J., Novakova R., Homeroova D., Rezuchova B.;  
 RT "Cloning and characterization of a new polypeptide gene cluster in  
 RT Streptomyces aureofaciens CCM3239."  
 RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AY033994; AAK61719.1; -  
 SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 84.3%; Score 43; DB 2; Length 420;  
 Best Local Similarity 100.0%; Pred. No. 1.7;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 4 GIRPVGRF 11  
 DB 48 GIRPVGRF 55  
 |||||

RESULT 3  
 ID 097950 PRELIMINARY; PRT; 426 AA.  
 AC 097950;  
 DT 01-OCT-2001 (TREMBlrel. 18, Created)  
 DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE TVG1120468 PROTEIN.  
 GN TVG1120468.  
 OS Thermoplasma volcanium.  
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;  
 OC Thermoplasma.  
 OX NCBI\_TaxID=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=SSSI / DSM 4299 / JCM 9571;  
 RX MEDLINE=20570466; PubMed=11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Kawashima O-hya Y., Watanabe K., Yamazaki K., Kanenori K., Kawamoto T.,  
 RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 RT sequence of Thermoplasma volcanium."  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 DR EMBL: AP000995; BAB60232.1; -  
 DR InterPro: IPR002881; DUF58.  
 DR InterPro: IPR002035; VMPA.  
 DR Pfam: PF01882; DUF58; 1.  
 DR SMART: SM00327; VMA; 1.  
 KM Complete Proteome.  
 SQ SEQUENCE 426 AA; 49092 MW; 63967CF711116F4A CRC64;

Query Match 70.6%; Score 36; DB 17; Length 426;  
 Best Local Similarity 75.0%; Pred. No. 46;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 4 GIRPVGRF 11  
 DB 304 GIRPVGRF 311  
 |||||

RESULT 4  
 ID 050431 PRELIMINARY; PRT; 674 AA.  
 AC 050431;  
 DT 01-JUN-1998 (TREMBlrel. 06, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)

DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE 2,4-DIENOYL-COA REDUCTASE.  
 GN FADH OR RV1175C OR MTW005.11C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 OX NCBI\_TaxID=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=H37RV;  
 RX MEDLINE=96295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 DR EMBL: AL010186; CAA15852.1; -  
 DR TubercuList: Rv1175c; -  
 DR InterPro: IPR001064; Crystallin.  
 DR InterPro: IPR001327; FAD\_pyr\_redox.  
 DR InterPro: IPR00205; NAD\_binding.  
 DR InterPro: IPR001155; Oxidored\_FMN.  
 DR Pfam: PF00724; oxidored\_FMN; 1.  
 DR Pfam: PF00070; pyr\_redox; 1.  
 DR PROSITE: PS00225; CRYSTALLIN\_BETACAMMA; UNKNOWN\_1.  
 KM Complete Proteome.  
 SQ SEQUENCE 674 AA; 72901 MW; 8F478F74ED3EE0A5 CRC64;

Query Match 70.6%; Score 36; DB 16; Length 674;  
 Best Local Similarity 85.7%; Pred. No. 75;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 4 GIRPVGR 10  
 DB 492 GIRPVGR 498  
 |||||

RESULT 5  
 ID 042924 PRELIMINARY; PRT; 272 AA.  
 AC 042924;  
 DT 01-JAN-1999 (TREMBlrel. 09, Created)  
 DT 01-JAN-1999 (TREMBlrel. 09, Last sequence update)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last annotation update)  
 DE HYPOTHEICAL 30.9 KDA PROTEIN C16A3.19 IN CHROMOSOME II.  
 GN SPEC16A3.19.  
 OS Schizosaccharomyces pombe (fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomyces.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=972;  
 RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Relhardt R.;  
 RL Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AL021748; CAA16870.1; -  
 DR InterPro: IPR001005; MYB\_DNA\_bind.  
 DR PROSITE: PS00037; MYB\_1; UNKNOWN\_1.  
 KM Hypothetical protein.  
 SQ SEQUENCE 272 AA; 30901 MW; 8FE0EB29217107A8 CRC64;

Query Match 68.6%; Score 35; DB 3; Length 272;  
 Best Local Similarity 85.7%; Pred. No. 45;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY      3  RGIRPVG 9
      11:1111
Db      47  RGLRPVG 53

RESULT  6
O981V3  PRELIMINARY; PRT; 371 AA.
AC      0981V3;
DT      01-OCT-2001 (TReMBLrel. 18, Created)
DT      01-OCT-2001 (TReMBLrel. 18, Last sequence update)
DE      MUR9222 PROTEIN.
GN      MUR9222.
OS      Rhizobium loti (Mesorhizobium loti).
OC      plasmid pMLA.
OC      Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC      Phyllobacteriaceae; Mesorhizobium.
OX      NCBI_TaxID=381;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=MAF303099;
RX      MEDLINE=21082930; PubMed=11214968;
RA      Kaneo T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA      Watanabe A., Iidesawa K., Ishikawa A., Kawashima K., Kimura T.,
RA      Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA      Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA      Takeuchi C., Yamada M., Tabata S.;
RT      "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT      Mesorhizobium loti."
RL      DNA Res. 7:331-338(2000).
DR      EMBL; AF003015; BAB54606.1; -.
KW      Plasmid; Complete proteome.
SQ      SEQUENCE 371 AA; 42385 MW; 98D9874B5E7831D1 CRC64;

Query Match      68.6%; Score 35; DB 16; Length 371;
Best Local Similarity 62.5%; Pred. No. 63;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY      3  RGIRPVG 10
      11:1111
Db      179  RGEPIGR 186

RESULT  7
O53895  PRELIMINARY; PRT; 504 AA.
AC      053895;
DT      01-JUN-1998 (TReMBLrel. 06, Created)
DT      01-JUN-1998 (TReMBLrel. 06, Last sequence update)
DT      01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE      PUTATIVE TWO-COMPONENT SENSOR.
GN      RV0992 OR MTV044.10.
OS      Mycobacterium tuberculosis.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1773;
RN      [1]
RP      SEQUENCE FROM N.A.
RC      STRAIN=H37RV;
RX      MEDLINE=98295987; PubMed=9634230;
RA      Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA      Gordon S.V., Eigemeier K., Gas S., Barry C.E. III, Tekala F.,
RA      Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA      Hornsby T., Jagels K., Kirogh A., McLean J., Moule S., Murphy L.,
RA      Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA      Rutter S., Seeger K., Skelton S., Squares R.,
RA      Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT      "Deciphering the biology of Mycobacterium tuberculosis from the
RT      complete genome sequence."

```

```

RL      Nature 393:537-544(1998).
CC      -1- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.
DR      EMBL; AL021999; CAI17581.1; -.
DR      HSSP; P02933; 1BXD.
DR      TubercuList; RV0982; -.
DR      InterPro; IPR003660; HAMF.
DR      InterPro; IPR003594; HATPase_C.
DR      InterPro; IPR003661; His_KinA.
DR      InterPro; IPR004359; HIS_KIN_s1g.
DR      Pfam; PF00672; HAMF; 1.
DR      Pfam; PF02518; HATPase_C; 1.
DR      Pfam; PF00512; signal; 1.
DR      SMART; SM00304; HAMF; 1.
DR      SMART; SM00387; HATPase_C; 1.
DR      SMART; SM00388; HisKA; 1.
KW      Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW      Transference
SQ      SEQUENCE 504 AA; 54407 MW; 26792A71AC432232 CRC64;

Query Match      68.6%; Score 35; DB 16; Length 504;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY      4  GIRPVG 10
      11:1111
Db      188  GLRPVGR 194

RESULT  8
O925G7  PRELIMINARY; PRT; 519 AA.
AC      0925G7;
DT      01-MAY-1999 (TReMBLrel. 10, Created)
DT      01-MAY-1999 (TReMBLrel. 10, Last sequence update)
DT      01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE      PUTATIVE TWO-COMPONENT SYSTEM SENSOR KINASE.
GN      ML0175 OR MLCB373.27.
OS      Mycobacterium lepre.
OC      Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC      Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX      NCBI_TaxID=1769;
RN      [1]
RP      SEQUENCE FROM N.A.
RA      Harris D., Taylor K.;
RL      Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN      [2]
RP      SEQUENCE FROM N.A.
RC      STRAIN=TN;
RX      MEDLINE=21128732; PubMed=11234002;
RA      Cole S.T., Eigemeier K., Parkhill J., James K.D., Thomson N.R.,
RA      Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA      Wungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA      Davies R.M., Devlin K., Dutfoy S., Feltham T., Fraser A., Hamlin N.,
RA      Holroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA      Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA      Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA      Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA      Barrell B.G.;
RT      "Massive gene decay in the leprosy bacillus."
RT      Nature 409:1007-1011(2001).
CC      -1- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.
DR      EMBL; AL035500; CAB36689.1; -.
DR      EMBL; AL583917; CAC29683.1; -.
DR      HSSP; P02933; 1BXD.
DR      LepToma; ML0175; -.
DR      InterPro; IPR004358; BCTRSENSOR.
DR      InterPro; IPR003660; HAMF.
DR      InterPro; IPR003594; HATPase_C.
DR      InterPro; IPR003661; His_KinA.
DR      InterPro; IPR004359; HIS_KIN_s1g.
DR      Pfam; PF00672; HAMF; 1.
DR      Pfam; PF02518; HATPase_C; 1.

```

DR Pfam: PF00512; signal: 1.  
 DR PRINTS: SM00344; BCTRLSENSOR.  
 DR SMART: SM00304; HAMF. 1.  
 DR SMART: SM00387; HATPase\_C. 1.  
 DR SMART: SM00388; HSKA. 1.  
 KW Complete proteome; kinase; phosphorylation; sensory transduction;  
 KW transference.  
 SQ SEQUENCE 519 AA; 56470 MW; C0F2DA4D57364F10 CRC64;

Query Match 68.6%; Score 35; DB 16; Length 519;  
 Best Local Similarity 85.7%; Pred. No. 90;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GRRPYGR 10  
 DB 188 GRRPYGR 194

RESULT 9  
 Q9KJ8 PRELIMINARY; PRT; 894 AA.  
 AC Q9KJ8; 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE DNA GYRASE, SUBUNIT A.  
 GN VC1258.  
 OS Vibrio cholerae.  
 OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
 OX NCBI\_TaxID=666;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
 RX MEDLINE=20406833; PubMed=10952301;  
 RA Heidelberg J.F., Eissen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
 RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
 RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
 RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,  
 RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
 RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
 RA Fraser C.M.;  
 RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
 cholerae";  
 RL Nature 406:447-483(2000).  
 DR EMBL: AE004205; AAF94417.1; -.  
 DR HSSP: P09097; IAB4.  
 DR TIGR: VC1258; -.  
 DR InterPro: IPR002205; DNA\_topoisomIV.  
 DR Pfam: PF00521; DNA\_topoisomIV. 1.  
 DR SMART: SM00434; TOP4C; 1.  
 KW Complete proteome.  
 SQ SEQUENCE 894 AA; 99025 MW; C7EA42056A08F57F CRC64;

Query Match 68.6%; Score 35; DB 16; Length 894;  
 Best Local Similarity 62.5%; Pred. No. 16+02;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 3 GRRPYGR 10  
 DB 733 GRRPYGR 740

RESULT 10  
 Q9VS40 PRELIMINARY; PRT; 1000 AA.  
 AC Q9VS40; 01-MAY-2000 (TREMblrel. 13, Created)  
 DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMblrel. 19, Last annotation update)  
 DE CG8608 PROTEIN.  
 GN UNC-13-4A OR CG8608.

OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abell J.F., Agbayani A., An H.-J., Andrews-Pfankuch C., Baldwin D.,  
 RA Ballieu R.M., Basu A., Baxendale J., Bayraktarglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Broxstein P., Brotlier P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Doup I.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferrier S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Kerchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Klup D., Lal Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Martel B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Paclet J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Slapson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003559; AAF50589.1; -.  
 DR HSSP: P21707; IBYN.  
 DR Flybase: FBgn0035756; unc-13-4A.  
 DR InterPro: IPR000008; C2.  
 DR Pfam: PF00168; C2; 1.  
 DR PRINTS: PR00360; C2DOMAIN.  
 DR SMART: SM00239; C2; 2.  
 DR PROSITE: PS00004; C2\_DOMAIN\_2; 1.  
 SQ SEQUENCE 1000 AA; 114286 MW; FDF14F05A55696A2 CRC64;

Query Match 68.6%; Score 35; DB 5; Length 1000;  
 Best Local Similarity 66.7%; Pred. No. 1.8+02;  
 Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 3 GRRPYGR 11  
 DB 130 GRRPYGR 138

RESULT 11  
 Q9PA00 PRELIMINARY; PRT; 1056 AA.  
 AC Q9PA00; 01-OCT-2000 (TREMblrel. 15, Created)  
 DT 01-OCT-2000 (TREMblrel. 15, Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)  
 DE HYPOTHEICAL PROTEIN XF2445.  
 GN XF2445.  
 OS Xylella fastidiosa.  
 OC Bacteria: Proteobacteria: gamma subdivision: Xanthomonas group;  
 OC Xylella.  
 NX NCBI\_TaxID=2371;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=9A5C;  
 RX MEDLINE=20365717; PubMed=10910347;  
 RA Simpson A.J.G., Retnach F.C., Arruda P., Abreu F.A., Acencio M.,  
 RA Alvarenga R., Alves L.M.C., Araya J.E., Baia G.S., Baptista C.S.,  
 RA Barrios M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Birones M.R.S.,  
 RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrer H.,  
 RA Colantu N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
 RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorry H.,  
 RA Pacinanti A.P., Ferreira A.J.S., Ferreira V.C.A., Fetto J.A.,  
 RA Fraga J.S., Franca S.C., Franco M.C., Froime M., Furlan L.R.,  
 RA Garner M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
 RA Ho P.L., Hohlsteil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
 RA Krieger J.E., Kuramae E.E., Laigret F., Lambis M.R., Leite L.C.C.,  
 RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
 RA Machado M.A., Madelira A.M.B.N., Martins E.M.F., Marino C.L.,  
 RA Marques M.A., Martins A.A.L., Martins E.M.F., Matsukuma A.Y.,  
 RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
 RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
 RA Nhami A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
 RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
 RA Peixoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,  
 RA Queiroz R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
 RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
 RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
 RA da Silveira J.F., Silveira M.L.Z., Silveira W.J., de Souza A.A.,  
 RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsunako M.H.,  
 RA Vallada H., Van Sluys M.A., Vertovaki-Almeida S., Vettore A.L.,  
 RA Zago M.A., Zatz M., Zeldanis J., Zetubal J.C.;  
 RT "The genome sequence of the plant pathogen Xylella fastidiosa.";  
 RL Nature 406:151-159(2000).  
 DR EMBL: AE004053; AAF85244.1; -;  
 DR InterPro: IPR000421; FAS58.C;  
 DR InterPro: IPR001092; HLH\_dtm.  
 DR Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 1056 AA; 118471 MW; 361CF2510BDFEC10 CRC64;

Query Match 68.6%; Score 35; DB 16; Length 1056;  
 Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 GIRPVGR 10  
 |:|||||  
 Db 1047 GLRPVGR 1053

RESULT 12  
 Q9VU79 PRELIMINARY; PRT; 1172 AA.  
 AC Q9VU79;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE CG10738 PROTEIN.  
 GN CG10738.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 NX NCBI\_TaxID=7227;  
 RN [1]  
 RP SROUENCE FROM N.A.  
 RC STRAIN=BRKLEY;  
 DE MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers J.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,  
 RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolhakov S.,  
 RA Borokova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Crawley S., Dahlke C., Davenport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K.J., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Fostler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris N.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Mostrel A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Munz D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
 RA Palazzolo M., Peltan G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,  
 RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zhang L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 DR EMBL: AE003537; AAF49811.1; -;  
 DR HSSP: 002846; IAWL.  
 DR FlyBase: FBgn0036368; CG10738.  
 DR InterPro: IPR001828; ANF\_receptor.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR InterPro: IPR001054; Guanylate\_cyclase.  
 DR Pfam: PF01094; ANF\_receptor\_2.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR SMART: SM00044; CYCC; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW ATP-binding; lyase; transferase.  
 SQ SEQUENCE 1172 AA; 131992 MW; C645093182A6D55 CRC64;

Query Match 68.6%; Score 35; DB 5; Length 1172;  
 Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 4 GIRPVGR 11  
 |:|||||  
 Db 434 GLRPVGR 441

RESULT 13  
 Q9RUR7 PRELIMINARY; PRT; 329 AA.  
 AC Q9RUR7;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE PUTATIVE ZINC-BINDING OXIDOREDUCTASE.  
 GN SCF51.18.

OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.J., Harris D.;  
 RL Submitted (Oct-1999) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Cerdeno A.M., Parthill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (Feb-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=9700351; PubMed=8643436;  
 RA Redenbach M., Kleser H.M., Denapalte D., Eichner A., Cullum J.,  
 RA Kinsahl H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL132707; CAB59716.1; -;  
 DR InterPro: IPR002085; Adh\_zinc\_family.  
 DR Pfam: PF00107; adh\_zinc.1.  
 SQ SEQUENCE 329 AA; 34155 MW; 5DCAD4FB174FD042 CRC64;

Query Match 66.7%; Score 34; DB 2; Length 329;  
 Best Local Similarity 87.5%; Pred. No. 88;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 3 RGIRPYGR 10  
 Db 137 RGIRYGR 144

RESULT 14  
 OQCLM8 PRELIMINARY; PRT; 529 AA.  
 AC 09CLM8:  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HEAT-INDUCED CATALASE.  
 OS Pleurotus sajor-caju (Oyster mushroom).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Agaricales; Pleurotaceae; Pleurotus.  
 OX NCBI\_TaxID=50053;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Jeong M.-J., Park S.-C.;  
 RT "Cloning of the catalase gene from Pleurotus sajor-caju.";  
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AF286097; AK15159.1; -;  
 DR HSP; P15202; 1A4E.  
 DR InterPro: IPR002226; Catalase.  
 DR Pfam: PF00199; catalase.1.  
 DR PRINTS: PR00067; CATALASE.  
 DR ProDom: PD000510; Catalase; 1.  
 DR PROSITE: PS00437; CATALASE\_1; 1.  
 DR PROSITE: PS00438; CATALASE\_2; UNKNOWN.1.  
 SQ SEQUENCE 529 AA; 59791 MW; A432BE6445B6AC13 CRC64;

Query Match 66.7%; Score 34; DB 3; Length 529;  
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 5 IRPVGR 11  
 Db 298 LRPIGRF 304

RESULT 15  
 ID 075230 PRELIMINARY; PRT; 537 AA.  
 AC 075230:  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE H53\_GSI (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Kimmerly W., Bonoc M., Cheng J., Connolly K.S., Gunning K.M.,  
 RA Kader K., Miguel T., Miller C., Plutuck S., Pollard M., Rojeski H.,  
 RA Subramanian S., Martin C.H.;  
 RT "Sequencing of human chromosome 5.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Ricke D.O.;  
 RT "Large Scale Sequence Analysis and Annotation with the Sequence  
 RT Comparison Analysis (SCAN) System.";  
 RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC005178; AAC24489.1; -;  
 FT NON\_TER 1  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 537 AA; 58017 MW; 2ED384E9DB64BCE5 CRC64;

Query Match 66.7%; Score 34; DB 4; Length 537;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 RGIRPYGR 10  
 Db 295 RGIRPFGK 302

Search completed: September 13, 2002, 09:29:23  
 Job time: 1068 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:36 : Search time 399.68 Seconds  
(without alignments)  
3.057 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_11\_21  
Perfect score: 50  
Sequence: 1 XRGIRPVGFRX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	10	20	AAW95178
2	48	96.0	15	18	AAW31400
3	48	96.0	15	20	AAW97230
4	48	96.0	15	21	AAI49293
5	48	96.0	15	21	AAI49295
6	48	96.0	15	21	AAI49296
7	48	96.0	20	18	AAW31394
8	48	96.0	20	18	AAW31387
9	48	96.0	20	18	AAW31374
10	48	96.0	20	20	AAW97232
11	48	96.0	20	20	AAW97234

12	48	96.0	20	20	AAW97236
13	48	96.0	20	20	AAW95191
14	48	96.0	20	20	AAW95175
15	48	96.0	20	21	AAI10350
16	48	96.0	20	21	AAI10358
17	48	96.0	20	21	AAI10365
18	48	96.0	20	21	AAI10369
19	48	96.0	20	21	AAI49294
20	48	96.0	20	21	AAI49301
21	48	96.0	20	21	AAI49302
22	48	96.0	20	22	AAI62519
23	48	96.0	20	22	AAI62527
24	48	96.0	20	22	AAI62534
25	48	96.0	20	22	AAI62538
26	48	96.0	20	22	AAI62539
27	48	96.0	20	22	AAI62540
28	48	96.0	20	22	AAI62541
29	48	96.0	20	22	AAI62542
30	48	96.0	20	22	AAI62543
31	48	96.0	20	22	AAI62544
32	48	96.0	20	22	AAI62545
33	48	96.0	20	22	AAI62546
34	48	96.0	20	22	AAI62547
35	48	96.0	20	22	AAI62548
36	48	96.0	20	22	AAI62549
37	48	96.0	20	22	AAI62550
38	48	96.0	20	22	AAI62551
39	48	96.0	20	22	AAI62552
40	48	96.0	20	22	AAI62553
41	48	96.0	20	22	AAI62554
42	48	96.0	20	22	AAI62555
43	48	96.0	20	22	AAI62556
44	48	96.0	20	22	AAI62557
45	48	96.0	20	22	AAI62558

## ALIGNMENTS

RESULT 1	
ID	AAW95178 standard; Protein: 10 AA.
AC	AAW95178:
XX	
XX	10-MAR-1999 (first entry)
DE	Murine pituitary-derived ligand polypeptide antigenic epitope.
XX	
KW	Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
KW	GR10; UHR-1; modulator; pituitary; central nervous system; pancreas;
KW	tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
KW	Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
KW	secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
KW	gene therapy; transgenic animal; epitope.
OS	Mus sp.
XX	
XX	WO9849295-A1.
XX	
PD	05-NOV-1998.
XX	
PF	27-APR-1998; 98WO-JP01923.
XX	
PR	28-APR-1997; 97JP-0109974.
XX	
PA	(TAKE ) TAKEDA CHEM IND LTD.
XX	
PI	Fukusumi S, Hinuma S;
XX	
DR	WPI; 1999-009423/01.
XX	

Human type ligand  
Bovine pituitary-d  
Murine pituitary-d  
Bovine oxytocin se  
Rat oxytocin secre  
Human oxytocin sec  
Oxytocin secretion  
19p2 ligand peptid  
19p2 ligand peptid  
19p2 ligand peptid  
Bovine CRH releas  
Rat CRH releas  
Human CRH releas  
CRH releasing prot  
Prolactin releasin  
Prolactin releasin  
Peptide PRP20 tra  
Synthetic G prote  
Human type G prote  
Rat type G prote  
Bovine G protein-c  
Partial ligand pol  
Mammalian 19p2 lig  
Bovine pituitary-d  
Bovine oxytocin se  
Rat oxytocin secre  
Human oxytocin sec  
Human CRH releas  
Rat CRH releas  
Human CRH releas  
Human type G prote  
Rat type G prote  
Bovine G protein-c

PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

XX Disclosure: Page 26; 206pp; English.

CC The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its nuclein are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.

XX Sequence 10 AA:

Query Match 96.0%; Score 48; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGR 10  
|||||  
DB 2 RGIIPVGR 10

#### RESULT 2

AAW31400  
ID AAW31400 standard; Peptide: 15 AA.

XX AAW31400;

DT 06-APR-1998 (first entry)

DE Synthetic ligand 19P2-L31 peptide II.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent; antigen.

XX Synthetic.

OS WO9724436-A2.

PN 10-JUL-1997.

PD 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

PA Fujii R, Fukusumi S, Habeta Y, Hinuma S, Hosoya M,

PI Kawamata Y, Kitada C;

XX WPI: 1997-363672/33.

DR Ligand peptide for G protein-coupled receptor - acts by modulating  
XX function in the central nervous system, pancreas and pituitary gland  
PT Example 43; Page 151; 258pp; English.

CC This peptide contains the partial C-terminal sequence of the synthetic  
CC ligand polypeptide 19P2-L31 which is capable of binding to a G  
CC protein-coupled receptor protein. This peptide is used as an antigen to  
CC prepare rabbit anti-bovine 19P2-L31 antibodies which are used in binding  
CC assays. Pharmaceutical compositions containing this ligand may be used  
CC as a pituitary function modulator, a central nervous system modulator  
CC or a pancreatic function modulator. This ligand could have specific  
CC applications as a prophylactic or therapeutic agent for dementia,  
CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
CC hyper- and hypophagia, hyperlipidaemia, hypercholesterolaemia,  
CC hyperglyceridaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
CC and/or oligogalactia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand  
CC affecting activation of the G protein-coupled receptor protein.

SO Sequence 15 AA:

Query Match 96.0%; Score 48; DB 18; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGR 10  
|||||  
DB 7 RGIIPVGR 15

#### RESULT 3

AAW97230  
ID AAW97230 standard; Peptide: 15 AA.

XX AAW97230;

DT 06-MAY-1999 (first entry)

DE C-terminal ligand polypeptide derived antigen.

XX G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;  
KW pituitary adenomatosis; brain tumour; emmenoptoly; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;

KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydatid mole;

KW interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocin; prolactin secretion.

XX Synthetic.

OS WO9858962-A1.

PN 30-DEC-1998.

PD 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PR (TAKE ) TAKEDA CHEM IND LTD.

PA Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

PI



XX WPI; 1999-105614/09.  
 DR  
 XX  
 PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Example 43; Page 123; 241pp; English.  
 PS  
 XX AAW97229-31 represent a ligand polypeptide derived fragments used to  
 CC produce antibodies. The specification describes an agent for modulating  
 CC prolactin secretion which comprises a ligand polypeptide or a salt, for  
 CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
 CC prolactin secretion can be used for treating or preventing  
 CC hypocoarctanism, gonocyst carcinogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,  
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,  
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly,  
 CC Chiari-Frommel syndrome, Argonzi-del Castillo syndrome, Forbes-Albright  
 CC syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory  
 CC agents can also be used as contraceptives. The agents for modulating  
 CC placental function can be used for treating or preventing  
 CC choriochorionoma, hydatid mole, abortion, unfertility  
 CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
 CC oxytocia.  
 CC  
 SO Sequence 15 AA;  
 XX  
 XX  
 Query Match 96.0%; Score 48; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0082;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 RGRPVGRF 10  
 Db 7 rgrlpvgrf 15  
 Db  
 XX  
 XX  
 RESULT 4  
 AAY49293  
 ID AAY49293 standard; peptide: 15 AA.  
 AC AAY49293;  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE 19P2 ligand peptide fragment.  
 XX  
 XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15  
 FT /note= "C-terminal amide"  
 XX  
 XX WO9960112-A1.  
 PN  
 XX  
 XX 25-NOV-1999.  
 PD  
 XX 20-MAY-1999; 99WO-JP02650.  
 PF  
 XX 21-MAY-1998; 98JP-0140293.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI  
 XX WPI; 2000-039381/03.  
 DR  
 XX

PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality .  
 XX  
 XX Disclosure; Page 26; 73pp; Japanese.  
 PS  
 XX The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.  
 CC  
 SO Sequence 15 AA;  
 XX  
 XX  
 Query Match 96.0%; Score 48; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.0082;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 2 RGRPVGRF 10  
 Db 7 rgrlpvgrf 15  
 Db  
 XX  
 XX  
 RESULT 5  
 AAY49295  
 ID AAY49295 standard; peptide: 15 AA.  
 AC AAY49295;  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE 19P2 ligand peptide fragment.  
 XX  
 XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 1  
 FT /note= "N-terminal acetylation"  
 FT Modified-site 4  
 FT /note= "acetylated Tyr"  
 FT Modified-site 15  
 FT /note= "C-terminal amide"  
 XX  
 XX WO9960112-A1.  
 PN  
 XX  
 XX 25-NOV-1999.  
 PD  
 XX 20-MAY-1999; 99WO-JP02650.  
 PF  
 XX 21-MAY-1998; 98JP-0140293.  
 PR  
 XX (TAKE ) TAKEDA CHEM IND LTD.  
 PA  
 XX Matsumoto H, Kitada C, Hinuma S;  
 PI  
 XX WPI; 2000-039381/03.  
 DR  
 XX  
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality .  
 XX  
 XX Disclosure; Page 26; 73pp; Japanese.  
 PS  
 XX The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function

CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA:

Query Match 96.0%; Score 48; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||||  
DB 7 rgirpvgrf 15

RESULT 6  
AAY49296  
ID AAY49296 standard; peptide: 15 AA.  
XX  
AC AAY49296;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE 19p2 ligand peptide fragment.

XX Monoclonal antibody: 19p2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
OS Homo sapiens.

XX  
FH Key Location/Qualifiers  
FT Modified-site 15  
FT /note= "C-terminal amide"  
XX

PM W09960112-A1.

PD 25-NOV-1999.

PF 20-MAY-1999; 99WO-JP02650.

PR 21-MAY-1998; 98JP-0140293.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI: 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
PS Disclosure: Page 27; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA:

Query Match 96.0%; Score 48; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||||  
DB 7 rgirpvgrf 15

RESULT 7  
AAW31394  
ID AAW31394 standard; peptide: 20 AA.  
XX  
AC AAW31394;  
XX  
DT 06-APR-1998 (first entry)  
XX

DE Human type G protein-coupled receptor ligand fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prolactin;  
KW therapeutic agent.  
XX  
OS Homo sapiens.

PN W09724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211803.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

DR WPI: 1997-363672/33.

DR N-PSDB; AAV02431.

PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland

PS Claim 2; Page 185; 258pp; English.

XX This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharide. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA:

Query Match 96.0%; Score 48; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10



RESULT 10  
 AAW97232 standard; peptide; 20 AA.  
 AC AAW97232;  
 DT 06-MAY-1999 (first entry)  
 DE Bovine pituitary-derived ligand polypeptide fragment.  
 KW Bovine pituitary-derived ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;  
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 OS Bos sp.  
 PN W09858962-A1.  
 PD 30-DEC-1998.  
 PF 22-JUN-1998; 98MO-JP02765.  
 PR 23-JUN-1997; 97JP-0165437.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 DR WPI: 1999-105614/09.  
 PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 secretion or placental function, e.g. for treating menopausal  
 syndrome, tumours, autoimmune disease or abnormal pregnancy  
 PS Claim 3; Page 136; 241pp: English.  
 CC The present sequence represents a bovine pituitary-derived ligand  
 fragment. It is used in the course of the invention. The specification  
 describes an agent for modulating prolactin secretion which comprises a  
 ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 protein. The agents for promoting prolactin secretion can be used for  
 treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
 syndrome, euthyroid or hypometabolism. They can be used for promoting  
 lactation in a domestic mammal and as an aphrodisiac. The agents for  
 inhibiting prolactin secretion can be used for treating or preventing  
 pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
 prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 modulating placental function can be used for treating or preventing  
 chorioncarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
 abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 SQ Sequence 20 AA;

Query Match 96.0%; Score 48; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIPIVGRF 10  
 |||||  
 DB 12 RGIPIVGRF 20

RESULT 11  
 AAW97234 standard; peptide; 20 AA.  
 AC AAW97234;  
 DT 06-MAY-1999 (first entry)  
 DE Rat type ligand polypeptide fragment.  
 KW Rat type ligand; modulation; prolactin secretion;  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;  
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia.  
 OS Rattus sp.  
 PN W09858962-A1.  
 PD 30-DEC-1998.  
 PF 22-JUN-1998; 98MO-JP02765.  
 PR 23-JUN-1997; 97JP-0165437.  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 DR WPI: 1999-105614/09.  
 PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 secretion or placental function, e.g. for treating menopausal  
 syndrome, tumours, autoimmune disease or abnormal pregnancy  
 PS Claim 3; Page 154; 241pp: English.  
 CC The present sequence represents a rat type ligand fragment. It  
 is used in the course of the invention. The specification describes  
 an agent for modulating prolactin secretion which comprises a  
 ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 protein. The agents for promoting prolactin secretion can be used for  
 treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
 syndrome, euthyroid or hypometabolism. They can be used for promoting  
 lactation in a domestic mammal and as an aphrodisiac. The agents for  
 inhibiting prolactin secretion can be used for treating or preventing  
 pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
 prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
 acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 modulating placental function can be used for treating or preventing  
 chorioncarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
 abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
 SQ Sequence 20 AA;

Query Match 96.0%; Score 48; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIPIVGRF 10  
 |||||  
 DB 12 RGIPIVGRF 20

RESULT 12  
ID AAM97236 standard; peptide: 20 AA.  
XX AAM97236:  
AC  
XX  
XX  
DT 06-MAY-1999 (first entry)  
XX  
XX  
DE Human type ligand polypeptide fragment.  
XX  
KM Rat type ligand: modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KM contraceptive; placental function; choriocarcinoma; hydralid mole;  
KM interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
KM abnormal lipidmetabolism; oxytocia.  
XX  
OS Homo sapiens.  
XX  
XX W09858962-A1.  
XX  
XX  
PD 30-DEC-1998.  
XX  
XX  
PF 22-JUN-1998: 98WO-JP02765.  
XX  
XX  
PR 23-JUN-1997: 97JP-0165437.  
XX  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX  
XX WPI: 1999-105614/09.  
XX  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
XX  
PS Claim 3: Page 166: 241pp: English.  
XX  
XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G-protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, amenorrhoea, galactorrhea,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydralid mole, interruption mole, abortion, unfertilized fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
XX  
SQ Sequence 20 AA:  
Query Match 96.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 13  
ID AAM95191 standard; peptide: 20 AA.  
XX AAM95191:  
XX  
XX  
AC  
XX  
XX  
DT 10-MAR-1999 (first entry)  
XX  
XX  
DE Bovine pituitary-derived ligand polypeptide fragment.  
XX  
XX  
KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; bovine.  
XX  
XX  
OS Bos sp.  
XX  
XX  
XX W09849295-A1.  
XX  
XX  
XX  
PD 05-NOV-1998.  
XX  
XX  
PF 27-APR-1998: 98WO-JP01923.  
XX  
XX  
PR 28-APR-1997: 97JP-0109974.  
XX  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX  
PI Fukusumi S, Hinuma S;  
XX  
XX WPI: 1999-009423/01.  
XX  
XX  
XX  
PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening  
XX  
XX  
PS Example 19: Page 151; 206pp: English.  
XX  
XX  
CC The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of the  
CC polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutant are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. The present sequence represents a bovine genome-derived ligand  
CC polypeptide fragment which is similar to the murine ligand-polypeptide.  
XX  
XX  
SQ Sequence 20 AA:  
Query Match 96.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 12 rgirpvgf 20

# RESULT 14

ID AAW95175 standard; Protein: 20 AA.

AAW95175;

10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
 KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
 KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
 KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
 KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
 KW gene therapy; transgenic animal; epitope.

OS Mus sp.

PN WO9849295-A1.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

DR WPI: 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used  
 for treating disorders of central nervous system, pituitary and  
 pancreas, and for drug screening

PS Disclosure: Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
 CC which is a ligand for the G-protein coupled orphan receptor designated  
 CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
 CC the ligand polypeptide encoding DNA are used to produce a recombinant  
 CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
 CC function of the pituitary, central nervous system, pancreas and other  
 CC tissues and can be used to screen for agents that modulate binding of  
 CC the polypeptide to the receptor; to quantify the amount of receptor in a  
 CC sample and to raise antibodies. They may also be used therapeutically,  
 CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
 CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
 CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
 CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
 CC operative nutritional status and as vasopressor. Transgenic animals  
 CC carrying the ligand polypeptide encoding DNA or its mutein are used to  
 CC study the function of the polypeptide-expressing genes, as models of  
 CC disease, for drug screening and as source of cell lines. The ligand  
 CC polypeptide DNA is used as a source of probes and primers; to identify  
 CC related sequences; in receptor-binding assays; for production of Ab and  
 CC antisera; in drug development; for gene therapy and to develop  
 CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
 CC epitopes which can be used for the preparation of anti-ligand polypeptide  
 CC antibody.

SQ Sequence 20 AA;

Query Match 96.0%; Score 48; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 2 RGIRPVGRF 10  
 |||||  
 Db 12 rgirpvgf 20

# RESULT 15

ID AAB10350 standard; peptide: 20 AA.

AAB10350;

24-NOV-2000 (first entry)

DE Bovine oxytocin secretion promoting peptide SEQ ID NO: 6.

KW Bovine; oxytocin secretion promoter; G protein-coupled receptor protein;  
 KW treatment; disease; pain; atonic bleeding; uterine recovery failure; cow;  
 KW caesarean section; artificial fertilization; galactostasis; goat; pig;  
 KW veterinary medicine; milk production.

OS Bos taurus.

PN WO200038704-A1.

PD 06-JUL-2000.

PF 22-DEC-1999; 99WO-JP07199.

PR 25-DEC-1998; 98JP-0369585.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;

DR WPI: 2000-452298/39.

PT Physiologically-active polypeptide recognized as ligand by G  
 protein-coupled receptor protein, for promoting secretion of oxytocin,  
 PR as drugs for diseases relating to oxytocin secretion and in veterinary  
 PT medicine

PS Claim 5; Page 51; 72pp; Japanese.

XX This invention describes a novel oxytocin secretion-regulating agent  
 CC which contains a ligand peptide or its salt for the G protein-coupled  
 CC receptor protein. It is useful in the form of drugs for ameliorating,  
 CC preventing and treating diseases relating to oxytocin secretion e.g.  
 CC weak pains and atonic bleeding, before and after expulsion of placenta,  
 CC uterine recovery failure, caesarean section, stoppage of artificial  
 CC fertilization or galactostasis and is also applicable in veterinary  
 CC medicine for promoting milk production in cow, goat and pig. This  
 CC sequence represents a bovine peptide which acts as an oxytocin secretion  
 CC promoter.

SQ Sequence 20 AA;

Query Match 96.0%; Score 48; DB 21; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.011;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RGIRPVGRF 10  
 |||||  
 Db 12 rgirpvgf 20

Search completed: September 13, 2002, 09:18:36  
 Job time: 501 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:58 ; Search time 136.62 Seconds  
(without alignments)  
1.967 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_11\_21  
Perfect score: 50  
Sequence: 1 XRGIRPVGRFX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued\_Patents\_AA: \*  
1: /cgn2\_6/prodata/2/1aa/5A.COMB.pep: \*  
2: /cgn2\_6/prodata/2/1aa/5B.COMB.pep: \*  
3: /cgn2\_6/prodata/2/1aa/6A.COMB.pep: \*  
4: /cgn2\_6/prodata/2/1aa/6B.COMB.pep: \*  
5: /cgn2\_6/prodata/2/1aa/PCTUS.COMB.pep: \*  
6: /cgn2\_6/prodata/2/1aa/Backfillsl.pep: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	15	4	US-08-776-971-93
2	48	96.0	20	3	US-09-105-678A-34
3	48	96.0	20	3	US-09-105-678A-40
4	48	96.0	20	3	US-09-105-678A-46
5	48	96.0	20	4	US-08-776-971-8
6	48	96.0	20	4	US-08-776-971-50
7	48	96.0	20	4	US-08-776-971-64
8	48	96.0	20	4	US-08-776-971-98
9	48	96.0	20	4	US-09-421-208-34
10	48	96.0	20	4	US-09-421-208-40
11	48	96.0	20	4	US-09-421-208-46
12	48	96.0	21	3	US-09-105-678A-28
13	48	96.0	21	3	US-09-105-678A-35
14	48	96.0	21	3	US-09-105-678A-41
15	48	96.0	21	3	US-09-105-678A-47
16	48	96.0	21	4	US-08-776-971-9
17	48	96.0	21	4	US-08-776-971-51
18	48	96.0	21	4	US-08-776-971-65
19	48	96.0	21	4	US-09-421-208-28
20	48	96.0	21	4	US-09-421-208-35
21	48	96.0	21	4	US-09-421-208-41
22	48	96.0	21	4	US-09-421-208-47
23	48	96.0	22	3	US-09-105-678A-36
24	48	96.0	22	3	US-09-105-678A-42
25	48	96.0	22	3	US-09-105-678A-48
26	48	96.0	22	4	US-08-776-971-10
27	48	96.0	22	4	US-08-776-971-52

28	48	96.0	22	4	US-08-776-971-66	Sequence 66, Appl
29	48	96.0	22	4	US-08-776-971-73	Sequence 73, Appl
30	48	96.0	22	4	US-09-421-208-36	Sequence 36, Appl
31	48	96.0	22	4	US-09-421-208-42	Sequence 42, Appl
32	48	96.0	22	4	US-09-421-208-48	Sequence 48, Appl
33	48	96.0	31	3	US-09-105-678A-7	Sequence 7, Appl
34	48	96.0	31	3	US-09-105-678A-8	Sequence 8, Appl
35	48	96.0	31	3	US-09-105-678A-9	Sequence 9, Appl
36	48	96.0	31	3	US-09-105-678A-31	Sequence 31, Appl
37	48	96.0	31	3	US-09-105-678A-37	Sequence 37, Appl
38	48	96.0	31	3	US-09-105-678A-43	Sequence 43, Appl
39	48	96.0	31	4	US-09-172-353-4	Sequence 4, Appl
40	48	96.0	31	4	US-08-776-971-5	Sequence 5, Appl
41	48	96.0	31	4	US-08-776-971-47	Sequence 47, Appl
42	48	96.0	31	4	US-08-776-971-61	Sequence 61, Appl
43	48	96.0	31	4	US-08-776-971-97	Sequence 97, Appl
44	48	96.0	31	4	US-09-421-208-7	Sequence 7, Appl
45	48	96.0	31	4	US-09-421-208-8	Sequence 8, Appl

## ALIGNMENTS

RESULT 1  
US-08-776-971-93  
Sequence 93, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47116  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-08-776-971-93

Query Match 96.0%; Score 48; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0024;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||||  
DB 7 RGIRPVGRF 15

RESULT 2  
US-09-105-678A-34  
; Sequence 34, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105.678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 96.0%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||||  
DB 12 RGIRPVGRF 20

RESULT 3  
US-09-105-678A-40  
; Sequence 40, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105.678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 96.0%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||||  
DB 12 RGIRPVGRF 20

RESULT 4  
US-09-105-678A-46  
; Sequence 46, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA



ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 96.0%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGRF 10  
Db 12 RGIRPVGRF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGRF 10  
Db 12 RGIRPVGRF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGRF 10  
|||||  
DB 12 RGIRPVGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGRF 10  
|||||  
DB 12 RGIRPVGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 98:  
US-08-776-971-98

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|1111111111|  
DB 12 RGIRPVGRF 20

RESULT 9  
US-09-421-208-34  
; Sequence 34, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10

DB 12 RGIRPVGRF 20  
|1111111111|

RESULT 10  
US-09-421-208-40  
; Sequence 40, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|1111111111|  
DB 12 RGIRPVGRF 20

RESULT 11  
US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 96.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGR 10  
DB 12 RGIRPVGR 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103862  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product= "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product= "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product= "Gly-OH or Gly-Arg"  
US-09-105-678A-28

Query Match 96.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGR 10  
DB 12 RGIRPVGR 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103862  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 96.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||||||  
DB 12 RGIRPVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 96.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 RGIRPVGRF 10  
|||||||

DB 12 RGIRPVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 96.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 2 RGIRPVGRF 10  
|||||||  
DB 12 RGIRPVGRF 20

Search completed: September 13, 2002, 09:20:59  
Job time: 624 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:02 ; Search time 172.41 Seconds  
(without alignments)  
6.131 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_11\_21  
Perfect score: 50  
Sequence: 1 XRGIRPVGRFX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues  
Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:.\*  
1: pir1:.\*  
2: pir2:.\*  
3: pir3:.\*  
4: pir4:.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	83	2 JC7607	prolactin-releasin
2	36	72.0	674	2 G70875	probable oxidoredu
3	35	70.0	272	2 T39533	hypothetical prote
4	35	70.0	391	1 XKG2AC	acetyl-coA C-acety
5	35	70.0	504	2 B70821	hypothetical prote
6	35	70.0	519	2 T45447	probable two-compo
7	35	70.0	894	2 E82221	DNA gyrase, chain
8	35	70.0	1056	2 B82557	hypothetical prote
9	34	68.0	226	2 A65057	hypothetical prote
10	34	68.0	307	2 S51485	ribosomal protein
11	34	68.0	328	2 S77236	ribosomal protein
12	34	68.0	343	2 AH1823	30S ribosomal prot
13	33	66.0	290	2 F97646	probable aliphatic
14	33	66.0	290	2 AD2870	ABC transporter, m
15	33	66.0	313	2 B95351	Vibrio type IV secr
16	33	66.0	319	2 T36845	probable membrane
17	33	66.0	335	2 E70855	hypothetical prote
18	33	66.0	353	2 C70562	hypothetical prote
19	33	66.0	366	2 C70365	twilching motility
20	33	66.0	445	2 T35893	FAD-dependent oxid
21	33	66.0	477	2 S68367	protoporphyrinogen
22	33	66.0	477	2 S65684	protoporphyrinogen
23	33	66.0	477	2 A56449	glutamyl-tRNA synt
24	33	66.0	485	2 E83653	glutamyl-tRNA synt
25	33	66.0	511	2 T40334	hypothetical prote
26	33	66.0	589	2 G87485	hypothetical prote
27	33	66.0	1940	2 F75393	hypothetical prote
28	32	64.0	57	2 I35055	MHC class II histo
29	32	64.0	149	2 AH2262	hypothetical prote

30	32	64.0	150	2 S09872	hypothetical prote
31	32	64.0	154	2 T34825	hypothetical prote
32	32	64.0	249	2 S76104	hypothetical prote
33	32	64.0	266	2 T43770	ribosomal protein
34	32	64.0	285	2 T50932	short-chain dehydr
35	32	64.0	282	2 G95926	probable saccharid
36	32	64.0	297	2 C70643	hypothetical prote
37	32	64.0	307	2 AD0883	conserved hypothet
38	32	64.0	309	2 T29118	hypothetical prote
39	32	64.0	374	2 G70947	hypothetical prote
40	32	64.0	388	2 AG3444	transporter BME115
41	32	64.0	409	2 T51126	hypothetical prote
42	32	64.0	410	2 D83573	tRNA nucleotidyl t
43	32	64.0	435	2 A72658	probable isocitrat
44	32	64.0	492	2 G70749	probable hycE prot
45	32	64.0	562	2 S16594	regulatory protein

## ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_rev150n 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; PMID:21092785; PMID:1178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PrRP  
A:Introns: 33/1

Query Match 96.0%; Score 48; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 RGIRPVGRF 10  
DB 44 RGIRPVGRF 52

RESULT 2  
G70875  
probable oxidoreductase (EC 1.5.99.-) - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_rev150n 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: G70875  
R:Cole, S.T.; Brosch, R.; Parhail, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, R.; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holtroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrall, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; PMID:9825987  
A:Accession: G70875  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-674 <COL>  
A:Cross-references: GB:AL010186; GB:AL123456; NID:93261493; PIDN:CA15852.1; PID:g269  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: fadH

C:Superfamily: Methylotrophus methylotrophus W3A1 trimethylamine dehydrogenase  
C:Keywords: 4Fe-4S; metalloprotein; oxidoreductase  
F:337,340,344,356/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 72.0%; Score 36; DB 2; Length 674;  
Best Local Similarity 85.7%; Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 GIRPVGR 9  
|:|||||  
DB 492 GIRPVGR 498

RESULT 3  
739533  
hypotheical protein SPBC16A3.19 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T39533  
R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.  
submitted to the EMBL Data Library, February 1998  
A:Reference number: 221862  
A:Accession: T39533  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-272 <MOO>  
A:Cross-references: EMBL:AL021748; PIDN:CAA16870.1; GSPDB:GN00067; SPDB:SPBC16A3.19  
C:Genetics:  
A:Gene: SPDB:SPBC16A3.19  
A:Map position: 2  
A:introns: 53/1; 97/3

Query Match 70.0%; Score 35; DB 2; Length 272;  
Best Local Similarity 85.7%; Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVG 8  
|:|||||  
DB 47 RGIRPVG 53

RESULT 4  
XGZAC  
acetyl-CoA C-acetyltransferase (EC 2.3.1.9) - Zoogloea ramigera  
N:Alternate names: acetoacetyl-CoA thiolase; biosynthetic thiolase; thiolase II  
C:Species: Zoogloea ramigera  
C:Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 05-May-2000  
C:Accession: A26121  
R:Peoples, O.P.; Masamune, S.; Walsh, C.T.; Sinskey, A.J.  
J. Biol. Chem. 262, 97-102, 1987  
A:Title: Biosynthetic thiolase from Zoogloea ramigera. III. Isolation and characterization  
A:Reference number: A26121; MUID:87083504  
A:Accession: A26121  
A:Molecule type: DNA  
A:Residues: 1-391 <PEO>  
A:Cross-references: EMBL:J02631; NID:q155617; PIDN:AAA27706.1; PID:q155618  
A:Experimental source: strain I-16-M, ATCC 19623  
C:Comment: The active enzyme, a tetramer of identical chains, catalyzes the reversible  
heats, and poly-beta-hydroxybutyrate biosynthesis.  
C:Genetics:  
A:Gene: phbA  
C:Superfamily: acetyl-CoA acetyltransferase  
C:Keywords: acyltransferase; coenzyme A; homotetramer; ketone body metabolism; poly-beta  
F:2-391/Product: acetyl-CoA acetyltransferase #status predicted <MAT>  
F:89/Active site: Cys #status predicted

Query Match 70.0%; Score 35; DB 1; Length 391;  
Best Local Similarity 75.0%; Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 RGIRPVGR 9  
|:|||||  
DB 266 RGIRPVGR 273

RESULT 5  
B70821  
hypotheical protein RV0982 - Mycobacterium tuberculosis (strain H37Rv)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70821  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno  
A:Reference number: A70500; MUID:98295987  
A:Accession: B70821  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-504 <COI>  
A:Cross-references: GB:AL021999; GB:AL123456; NID:q3261538; PIDN:CAA17581.1; PID:q291  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV0982  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 70.0%; Score 35; DB 2; Length 504;  
Best Local Similarity 85.7%; Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 GIRPVGR 9  
|:|||||  
DB 188 GIRPVGR 194

RESULT 6  
T45447  
probable two-component system sensor [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C:Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T45447  
R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, February 1998  
A:Reference number: 222967  
A:Accession: T45447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-519 <JAM>  
A:Cross-references: EMBL:AL035500; PIDN:CAB36689.1  
A:Experimental source: cosmid L573  
C:Genetics:  
A:Note: MLCB373.27  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 70.0%; Score 35; DB 2; Length 519;  
Best Local Similarity 85.7%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 GIRPVGR 9  
|:|||||  
DB 188 GIRPVGR 194

RESULT 7  
B82221  
DNA gyrase, chain A VC1258 [imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001



C:Accession: E82221  
R.Heldberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwin, M.L.; Dodson, R.J.;  
Chaidson, D.; Ermolaeva, M.D.; Vamathavan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, H.  
1, R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: AB2035; MUID:20406833  
A:Accession: E82221  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-894 <HEI>  
A:Cross-references: GB:AE004205; GB:AE003852; NID:g9555740; PIDN:AAF94417.1; GSPDB:GN001  
A:Experimental source: serogroup O1; strain N16961; biotype El Tor  
C:Genetics:  
A:Gene: VC1258  
A:Map position: 1  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase C

Query Match 70.0%; Score 35; DB 2; Length 894;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;  
QY 2 RGIRPYGR 9  
DB 733 KGVPRMGR 740

RESULT 8  
B82557  
hypothetical protein XF2445 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82557  
Ranomymus, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: AB2515; MUID:20365717  
A:Note: For a complete list of authors see reference number A59328 below  
A:Accession: B82557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1056 <SIM>  
A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF85244.1; GSPDB:GN001  
A:Experimental source: strain 9a5c  
R:Simpton, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alvarenga, R.; A  
Brlones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrier, H  
as-Neto, E.; Docena, C.; El-Doory, H.; Facincani, A.P.; Ferreira, A.J.S.  
submitted to GenBank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; Laig  
chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Martino, C.L.; Marques, M.V.; Martins, H  
A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmeri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silveir  
M.; Tshahko, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovsky-Almeida, S.; Vettore, A.L.; Z  
A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2445

Query Match 70.0%; Score 35; DB 2; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 92;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 3 GTRPVGR 9  
DB 1047 GTRPVGR 1053

RESULT 9

A65057  
hypothetical protein b2757 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C:Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
C:Accession: A65057  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: A65057  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-226 <BLAT>  
A:Cross-references: GB:AE000359; GB:U00096; NID:g1789110; PIDN:AACT5799.1; PID:g17891  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypothetical protein b2757

Query Match 68.0%; Score 34; DB 2; Length 226;  
Best Local Similarity 75.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
QY 3 GTRPVGR 10  
DB 25 GTRPVGR 32

RESULT 10  
S51485  
ribosomal protein S1 - Synechococcus sp. (PCC 6301)  
C:Species: Synechococcus sp.  
A:Variety: PCC 6301  
C:Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000  
C:Accession: S51485  
R:Sugita, M.; Sugita, C.; Sugita, M.  
Mol. Gen. Genet. 246, 142-147, 1995  
A:Title: Structure and expression of the gene encoding ribosomal protein S1 from the  
n CSI.  
A:Reference number: S51483; MUID:95166170  
A:Accession: S51485  
A:Molecule type: DNA  
A:Residues: 1-307 <SUG>  
A:Cross-references: EMBL:D28752; NID:g560122; PIDN:BA05946.1; PID:g666973  
A:Experimental source: PCC 6301  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: Synechocystis ribosomal protein S1  
C:Keywords: protein biosynthesis; ribosome

Query Match 68.0%; Score 34; DB 2; Length 307;  
Best Local Similarity 66.7%; Pred. No. 41;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
QY 2 RGIRPYGR 10  
DB 205 RGIRPYGR 213

RESULT 11  
S77236  
ribosomal protein S1 - Synechocystis sp. (strain PCC 6803)  
N:Alternate names: protein slr1356  
C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C:Accession: S77236  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asanizu, E.; Nakamura, Y.; Miyajima,  
O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.

A:Reference number: S74322; MUID:97061201  
A:Accession: S77236  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-328 <KAN>  
A:Cross-references: EMBL:D90907; GB:AB001339; NID:g1652618; PIDN:BA017570.1; PID:g165265  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: Synecocystis ribosomal protein S1  
C:Keywords: protein biosynthesis; ribosome

Query Match 68.0%; Score 34; DB 2; Length 328;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||:| |  
Db 204 RGIRPVGRF 212

RESULT 12  
AH1823  
30S ribosomal protein S1 [Imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH1823  
R:Kaneke, T.; Nakamura, Y.; Wolk, C.P.; Kurtz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH1823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-343 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA077660.1; PID:g17135114; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: Synecocystis ribosomal protein S1

Query Match 68.0%; Score 34; DB 2; Length 343;  
Best Local Similarity 66.7%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||:| |  
Db 205 RGIRPVGRF 213

RESULT 13  
F97646  
Probable aliphatic sulfonates transport permease protein ssuc AGR\_C\_4335 [Imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: F97646  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Mollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; PMID:11743194  
A:Accession: F97646  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AK88127.1; PID:g15157563; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C\_4335  
A:Map position: circular chromosome

Query Match 66.0%; Score 33; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||:| |  
Db 123 RGLPPIGRF 131

RESULT 14  
AD2870  
ABC transporter, membrane spanning protein Atu2390 [Imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AD2870  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woorag, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McCl ; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AD2870  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AA43378.1; PID:g17740875; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu2390  
A:Map position: circular chromosome

Query Match 66.0%; Score 33; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
|||:| |  
Db 123 RGLPPIGRF 131

RESULT 15  
B95351  
VirB type IV secretion protein [Imported] - Sinorhizobium meliloti (strain 1021) mag  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: B95351  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B ; Kallman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: B95351  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AA65372.1; PID:g14523833; GSPDB:GN00165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Galibert, F.; Flman, T.M.; Long, S.R.; Pulver, A.; Abola, P.; Ampe, F.; Barloy-Hubler, D.; Chai, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kallman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelauheault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
C:Contents: annotation  
C:Genetics:  
A:Gene: VirB6

A:Genome: plasmid

Query Match 66.0%; Score 33; DB 2; Length 313;  
Best Local Similarity 75.0%; Pred. No. 66;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 2 RGI RPYGR 9  
||:| |||  
Db 297 RGVRAVGR 304

Search completed: September 13, 2002, 09:24:03  
Job time: 778 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:48 ; Search time 80.21 Seconds  
(without alignments)  
5.310 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_11\_21  
Perfect score: 50  
Sequence: 1 XRGIRPVGRFX 11

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	96.0	83	1 PRRP_RAT	P81278 ratu8 norv
2	48	96.0	87	1 PRRP_HUMAN	P81277 homo sapien
3	48	96.0	98	1 PRRP_BOVIN	P81264 bos taurus
4	35	70.0	391	1 THIL_ZOORA	P07097 zoogloea ra
5	35	70.0	922	1 GYRA_AERSA	P48369 aeromonas s
6	34	68.0	224	1 YGCI_ECOLI	P46398 escherichia
7	34	68.0	306	1 RSL_STYNG	P46228 synchococc
8	34	68.0	328	1 RSL_STYNG	P73330 synchococc
9	34	68.0	801	1 TFR2_HUMAN	P90552 homo sapien
10	33	66.0	477	1 PPOX_HUMAN	P50336 homo sapien
11	33	66.0	477	1 PPOX_HUMAN	P50336 homo sapien
12	33	66.0	485	1 SYE_BACHD	P09466 bacillus ha
13	32	64.0	150	1 ULAY_HCMVA	P16828 human cytom
14	32	64.0	249	1 Y361_SYNY3	O55578 synchocyst
15	32	64.0	330	1 AR72_HUMAN	O43488 homo sapien
16	32	64.0	331	1 AR73_HUMAN	O93154 homo sapien
17	32	64.0	676	1 EXL1_HUMAN	O92935 homo sapien
18	31	62.0	130	1 RL3_PIG	O29293 sus scrofa
19	31	62.0	149	1 ENRM_BPT7	P00641 bacterioph
20	31	62.0	207	1 PYR5_SCHPO	P44198 schizosacch
21	31	62.0	230	1 PYR5_SCHPO	P44198 schizosacch
22	31	62.0	233	1 PYR5_SCHPO	P44198 schizosacch
23	31	62.0	250	1 LINC_PSEPA	P50197 pseudomonas
24	31	62.0	364	1 MK11_MOUSE	O9W11 mus musculu
25	31	62.0	367	1 GCST_MYCTU	O10376 mycobacteri
26	31	62.0	372	1 MK11_HUMAN	O15759 homo sapien
27	31	62.0	413	1 CSD_THEMA	O9X191 thermotoga
28	31	62.0	430	1 PYR5_MYCTU	P18009 mycobacteri
29	31	62.0	433	1 ZUOI_YEAST	P32527 saccharomyc
30	31	62.0	463	1 K10_DROME	P13468 drosophila
31	31	62.0	625	1 DHGL_DROME	P18173 drosophila
32	31	62.0	625	1 DHGL_DROME	P18173 drosophila
33	31	62.0	631	1 DNAM_THERO	P96133 thermomicro

34	31	62.0	719	1 NRPI_YEAST	P32770 saccharomyc
35	31	62.0	908	1 ACON_BACSU	P09339 bacillus su
36	31	62.0	1056	1 POL_BIV06	P19560 bovine immu
37	31	62.0	1056	1 POL_BIV27	P19561 bovine immu
38	31	62.0	1416	1 BLM_MOUSE	O68700 mus musculu
39	31	62.0	1561	1 SPAD_STRMU	P23504 streptococc
40	31	62.0	1565	1 PAC_STRMU	P11557 streptococc
41	30	60.0	116	1 RL34_HUMAN	P49207 homo sapien
42	30	60.0	179	1 RK6_GUTTH	O46508 guillardi
43	30	60.0	292	1 CNTB_HUMAN	O9UFF9 homo sapien
44	30	60.0	368	1 YG3Y_YEAST	P53295 saccharomyc
45	30	60.0	376	1 MPK5_ARATH	O39025 arabidopsis

## ALIGNMENTS

RESULT 1  
PRRP\_RAT STANDARD; PRT; 83 AA.  
AC P81278;  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Proactin-releasing peptide precursor (Prp) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide Prp31; Proactin-releasing peptide Prp20].  
DE PRH.  
GN Rattus norvegicus (Rat).  
OS Eukaryota; Metazoa; Chordata; Cranata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus. OX NCBI\_TaxID=10116;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=98268781; PubMed=9607765;  
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
RT "A proactin-releasing peptide in the brain.";  
RL Nature 393:272-276(1998).  
RN [2]  
RP TISSUE SPECIFICITY.  
RX PubMed=10498338;  
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
RT "Tissue distribution of proactin-releasing peptide (Prp) and its receptor.";  
RL Regul. Pept. 83:1-10(1999).  
CC -!- FUNCTION: Stimulates proactin (Prp) release and regulates the expression of proactin through its receptor GPR10. May stimulate lactotrophs directly to secrete Prp.  
CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in medulla oblongata and hypothalamus.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC EMBL AB015418; BAA29026.1; -  
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
FT SIGNAL 1 21  
FT PEPTIDE 22 52 PROACTIN-RELEASING PEPTIDE PRP31.  
FT PEPTIDE 33 52 PROACTIN-RELEASING PEPTIDE PRP20.  
FT MOD\_RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 83 AA; 9215 MW; DDC75A264EEBAF29 CXC64;

Query Match 96.0%; Score 48; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.004;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RCIRPYGRF 10  
 |||||  
 DB 44 RCIRPYGRF 52

# RESULT 2 PRRP\_HUMAN

ID PRRP\_HUMAN STANDARD; PRT; 87 AA.

AC P81277; 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP3; Proactin-releasing peptide PRRP2].  
 DE PRRP.  
 GN Homo sapiens (Human).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 OC NCBI\_TaxID=9606;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP TISSUE=Brain;  
 RC MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada S., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A proactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 RN [2]  
 RP TISSUE SPECIFICITY.  
 RX PubMed=10498338;  
 RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RA "Tissue distribution of proactin-releasing peptide (PRRP) and its receptor.";  
 RT Regul. Pept. 83:1-10(1999).  
 RL -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL; AB015419; BAA29027.1; -  
 DR MIM; 602663; -  
 KW Hormone; Amidation; Signal.  
 FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP3.  
 FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

Query Match 96.0%; Score 48; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RCIRPYGRF 10  
 |||||  
 DB 45 RCIRPYGRF 53

RESULT 3  
 PRRP\_BOVIN  
 ID PRRP\_BOVIN STANDARD; PRT; 98 AA.

AC P81264; 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP3; Proactin-releasing peptide PRRP2].  
 DE PRRP.  
 GN Bos taurus (Bovine).  
 OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; Bovidae; Bos.  
 OC NCBI\_TaxID=9913;  
 OX [1]  
 RN SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.  
 RP TISSUE=Brain;  
 RC MEDLINE=98268781; PubMed=9607765;  
 RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada S., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RT "A proactin-releasing peptide in the brain.";  
 RL Nature 393:272-276(1998).  
 RN -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sdb.ch/announce/> or send an email to [license@isb-sdb.ch](mailto:license@isb-sdb.ch)).  
 CC -----  
 DR EMBL; AB015417; BAA29025.1; -  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.  
 FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROLACTIN-RELEASING PEPTIDE PRRP3.  
 FT PEPTIDE 33 53 PROLACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 96.0%; Score 48; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RCIRPYGRF 10  
 |||||  
 DB 45 RCIRPYGRF 53

# RESULT 4 THIL\_ZOORA

ID THIL\_ZOORA STANDARD; PRT; 391 AA.

AC P07097; 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Acetyl-CoA acetyltransferase (EC 2.3.1.9) (Acetoacetyl-CoA thiolase).  
 DE PHA.  
 GN Zoogloa ramigera.  
 OS Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group; Zoogloa.  
 OC NCBI\_TaxID=350;  
 OX [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=ATCC 19623 / I-16-M;

```

RX MEDLINE-87083504; PubMed-2878929;
RA Peoples O.P., Masamune S., Walsh C.T., Sinskey A.J.;
RT "Biosynthetic thiolase from Zoogloa ramifera. III. Isolation and
RT characterization of the structural gene.";
RL J. Biol. Chem. 262:97-102(1987).
RN [2]
RN REVISON TO 130.
RX MEDLINE-89359356; PubMed-2670935;
RA Peoples O.P., Sinskey A.J.;
RT "Poly-beta-hydroxybutyrate biosynthesis in Alcaligenes eutrophus H16.
RT characterization of the genes encoding beta-ketothiolase and
RT acetyl-CoA reductase.";
RL J. Biol. Chem. 264:15293-15297(1989).
RN [3]
RN MUTAGENESIS OF CYS-377.
RX MEDLINE-91217075; PubMed-1673680;
RA Palmer M.A.J., Differding E., Gamboni R., Williams S.F., Peoples O.P.,
RA Walsh C.T., Sinskey S.J., Masamune S.;
RT Biosynthetic thiolase from Zoogloa ramifera. Evidence for a
RT mechanism involving Cys-378 as the active site base.";
RL J. Biol. Chem. 266:8369-8375(1991).
CC -1- CATALYTIC ACTIVITY: 2 acetyl-CoA = CoA + acetyl-CoA.
CC -1- PATHWAY: FIRST STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J02631; AAA27706.1; ALT_SEQ.
DR PIR: A26121; XKGZAC.
DR PIR: A27754; A27754.
DR HSSP: P27796; IAFY.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase_1.
DR PROSITE: PS00098; THIOLASE_1.
DR PROSITE: PS00099; THIOLASE_3.
DR PROSITE: PS00737; THIOLASE_2.
DR Transferrase; Acyltransferase; PHB biosynthesis.
FT INIT MET 0
FT ACT_SITE 88 88 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 377 377 BASE.
FT MUTAGEN 377 377 C->G: LOSS OF ACTIVITY.
SQ SEQUENCE 391 AA; 40342 MW; 6D2351A1BC0E4EDD CRC64;

```

Query Match 70.0%; Score 35; DB 1; Length 391;  
 Best Local Similarity 75.0%; Pred. No. 9.3;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 RGIRPVR 9
DB 266 RGIRPVR 273

```

RESULT 5  
 ID GYRA\_AERSA STANDARD; PRT; 922 AA.  
 AC P48369;  
 DT 01-FEB-1996 (Rel. 33, Created)  
 DT 01-FEB-1996 (Rel. 33, Last sequence update)  
 DT 01-NOV-1997 (Rel. 35, Last annotation update)  
 DE DNA gyrase subunit A (EC 5.99.1.3).  
 GN GYRA.  
 OS Aeromonas salmonicida.  
 OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;

```

OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RN SEQUENCE FROM N.A.
RP STRAIN=2148/89;
RA Oppergaard H.;
RL Submitted (OCT-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RN SEQUENCE OF 33-179 FROM N.A.
RP STRAIN=ATCC 14174;
RX MEDLINE-95142596; PubMed-7840589;
RA Oppergaard H., Sorum H.;
RT "gyrA mutations in quinolone-resistant isolates of the fish pathogen
RT Aeromonas salmonicida.";
RL Antimicrob. Agents Chemother. 38:2460-2464(1994).
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING. THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: I47978; AAB41037.1; -.
DR EMBL: I42453; AAA87239.1; -.
DR HSSP: P09097; IAB4.
DR InterPro: IPR002205; DNA_topoisomIV.
DR Pfam: PF00521; DNA_topoisomIV.
DR SMART: SM00434; TOP4c.
DR Isomerase; Topoisomerase; DNA-binding.
FT ACT_SITE 122 122 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 922 AA; 101333 MW; 8894965DC4217077 CRC64;

```

Query Match 70.0%; Score 35; DB 1; Length 922;  
 Best Local Similarity 62.5%; Pred. No. 22;  
 Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

OY 2 RGIRPVR 9
DB 758 KGVRPVR 765

```

RESULT 6  
 ID YGCI\_ECOLI STANDARD; PRT; 224 AA.  
 AC Q46898;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein ygcI precursor.  
 GN YGCI OR B2757.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RN SEQUENCE FROM N.A.  
 RP STRAIN=K12 / MG1655;  
 RX MEDLINE-97426617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Berna N.T., Burland V.,  
 RA Riley M., Coliado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

RA Mau B., Shao Y.;  
RT "The complete genome sequence of Escherichia coli K-12."  
RL Science 277:1453-1474(1997).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: U29579; AAA69267.1; ALT\_INIT.  
DR EMBL: AE000359; AAC75799.1; ALT\_INIT.  
DR Ecogen: BG13116; ygcI.  
KW Hypothetical protein; Signal; Complete proteome.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 224 HYPOTHETICAL PROTEIN YGC1.  
SQ SEQUENCE 224 AA: 25209 MW: 1C42CC09B317D68 CRC64;  
  
Query Match 68.0%; Score 34; DB 1; Length 224;  
Best Local Similarity 75.0%; Pred. No. 8.5;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
OY 3 GIRPYGRF 10  
DB 23 GRRPYGRF 30  
  
RESULT 7  
RS1\_SYNP6 STANDARD; PRT; 306 AA.  
ID RS1\_SYNP6  
AC P46228;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 30S ribosomal protein S1.  
DE RPSA OR RPS1.  
OS Synecococcus sp. (strain PCC 6301) (Anacystis nidulans).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecococcus.  
OX NCBI\_TaxID=1139;  
RN [1]  
RP SEQUENCE FROM N.A., AND SEQUENCE OF 1-13.  
RA MEDLINE=9516170; PubMed=7662084;  
RA Sugita M., Sugita C., Sugitara M.;  
RT "Structure and expression of the gene encoding ribosomal protein S1  
RT from the cyanobacterium Synecococcus sp. strain PCC 6301: striking  
RT sequence similarity to the chloroplast ribosomal protein CSI.";  
RL Mol. Gen. Genet. 246:142-147(1995).  
RL -1- FUNCTION: BINDS MRNA.  
CC -1- SIMILARITY: BELONGS TO THE SLP FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 3 S1 MOTIF DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: D28752; BAA05946.1; -.  
DR HSSP: P05055; ISRO.  
DR InterPro: IPR003029; S1.  
DR Pfam: PF00575; S1; 3.  
DR SMART: SM00316; S1; 3.  
DR PROSITE: PS50126; S1; 3.  
KW Ribosomal protein; Repeat; RNA-binding.  
FT INIT\_MET 0  
FT DOMAIN 31 100 S1 MOTIF 1.  
FT DOMAIN 118 182 S1 MOTIF 2.  
FT DOMAIN 196 264 S1 MOTIF 3.

SQ SEQUENCE 306 AA: 34492 MW: 82D4A2EBE69B97CE CRC64;  
  
Query Match 68.0%; Score 34; DB 1; Length 306;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
OY 2 RGIRPYGRF 10  
DB 204 RGIRPYGARF 212  
  
RESULT 8  
RS1A\_SYNY3 STANDARD; PRT; 328 AA.  
ID RS1A\_SYNY3  
AC P73530;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 30S ribosomal protein S1 homolog A.  
GN RPS1A OR SR1356.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905221;  
RA Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naito K.,  
RA Okumura S., Shimpō S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
RL -1- FUNCTION: BINDS MRNA.  
CC -1- SIMILARITY: BELONGS TO THE SLP FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 3 S1 MOTIF DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL: D90907; BAA17570.1; -.  
DR HSSP: P05055; ISRO.  
DR InterPro: IPR003029; S1.  
DR Pfam: PF00575; S1; 3.  
DR SMART: SM00316; S1; 3.  
DR PROSITE: PS50126; S1; 3.  
KW Ribosomal protein; Repeat; RNA-binding; Complete proteome.  
FT DOMAIN 31 100 S1 MOTIF 1.  
FT DOMAIN 118 182 S1 MOTIF 2.  
FT DOMAIN 196 264 S1 MOTIF 3.  
SQ SEQUENCE 328 AA: 36570 MW: DC3F9B1E5A40619 CRC64;  
  
Query Match 68.0%; Score 34; DB 1; Length 328;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
OY 2 RGIRPYGRF 10  
DB 204 RGIRPYGARF 212  
  
RESULT 9  
TFR2\_HUMAN STANDARD; PRT; 801 AA.  
ID TFR2\_HUMAN



AC Q9UP53: 075422; Q9HA99; C9AK67  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Transferrin receptor protein 2 (TfR2).  
GN TfR2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORMS ALPHA AND BETA).  
RP TISSUE=Erythroleukemia, and Myeloid leukemia cells;  
RC MEDLINE=99340005; PubMed=10409623;  
RA Kawabata H., Yang R., Hirama T., Vuong P.T., Kawano S., Gombart A.F.,  
RA Koefler H.P.,  
RT "Molecular cloning of transferrin receptor 2: a new member of the  
RT transferrin receptor-like family."  
RL J. Biol. Chem. 274:20826-20832(1999).  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM GAMMA).  
RX PubMed=9799793;  
RA Gloeckner G., Scherer S., Schatterov R., Borlight A.P., Weber J.,  
RA Tsui L.-C., Rosenthal A.,  
RT "Large scale analysis of two regions in human chromosome 7q22:  
RT annotation of 650 kb of genomic sequence around the EPO and CNTF1 loci  
RT reveals 17 genes."  
RL Genome Res. 8:1060-1073(1998).  
RN [3]  
RP SEQUENCE OF 1-158 AND 370-801 FROM N.A.  
RC TISSUE=Carcinoma, and Embryo;  
RA Isegal T., Oca T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Nishikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Wagatsuma M., Hosokari T., Kaku Y., Kodaira H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Nakakawa K., Ono Y., Takiguchi S.,  
RA Watanabe S., Kimura K., Murakami K., Ishi S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu K., Nakamura Y., Nagahata K., Masuno Y.,  
RA Niinomiya K., Iwayanagi T.,  
RT "NEDO human cDNA sequencing project."  
RT Submitted (Aug-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RP DISEASE.  
RX PubMed=10802645;  
RA Camaschella C., Roetto A., Call A., De Gobbi M., Garozzo G.,  
RA Carella M., Majofano N., Totaro A., Gasparini P.,  
RT "The gene TFR2 is mutated in a new type of haemochromatosis mapping to  
RT 7q22."  
RL Nat. Genet. 25:14-15(2000).  
RN [5]  
RP VARIANT: HFE3 LYS-172.  
RX PubMed=11313241;  
RA Roetto A., Totaro A., Piperno A., Piga A., Longo F., Garozzo G.,  
RA Call A., De Gobbi M., Gasparini P., Camaschella C.,  
RT "New mutations inactivating transferrin 2 in hemochromatosis type  
RT 3."  
RL Blood 97:2555-2560(2001).  
RN [6]  
RP FUNCTION: Mediates cellular uptake of transferrin-bound iron in a  
CC non-iron dependent manner. May be involved in iron metabolism,  
CC hepatocyte function and erythrocyte differentiation.  
RN [7]  
RP SUBUNIT: Homodimer.  
CC [8]  
RP SUBCELLULAR LOCATION: Type II membrane protein. The beta isoform  
CC lacks the transmembrane domain and is probably intracellular.  
RN [9]  
RP ALTERNATIVE PRODUCTS: 3 isoforms: alpha (shown here), beta and  
CC gamma: are produced by alternative splicing.  
RN [10]  
RP TISSUE SPECIFICITY: Predominantly expressed in liver. While the  
CC alpha form is also expressed in spleen, lung, muscle, prostate and  
CC peripheral blood mononuclear cells, the beta form is expressed in  
CC all tissues tested, albeit weakly.  
RN [11]  
RP DISEASE: Defects in TFR2 are a cause of hereditary hemochromatosis  
CC type III (HFE3). HFE3 is a disorder of iron homeostasis resulting  
CC in iron overload and has a phenotype indistinguishable from that  
CC of hereditary hemochromatosis (HH). HH is characterized by  
CC abnormal intestinal iron absorption and progressive increase of

CC		totalbody iron, which results in midlife in clinical
CC		complications including cirrhosis, cardiopathy, diabetes,
CC		endocrine dysfunctions, arthropathy, and susceptibility to liver
CC		cancer. Since the disease complications can be effectively
CC		prevented by regular phlebotomies, early diagnosis is most
CC		important to provide a normal life expectancy to the affected
CC		subjects.
CC	-1-	MISCELLANEOUS: The variant lys-172 found in hereditary
CC		hemochromatosis type III affects the putative initiation codon of
CC		the beta isoform thus preventing its translation.
CC	-1-	SIMILARITY: BELONGS TO PEPTIDASE FAMILY M28B.
CC		-----
CC		This SWISS-PROT entry is copyright. It is produced through a collaboration
CC		between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC		the European Bioinformatics institute. There are no restrictions on its
CC		use by non-profit institutions as long as its content is in no way
CC		modified and this statement is not removed. Usage by and for commercial
CC		entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>
CC		or send an email to <a href="mailto:license@isb-ib.ch">license@isb-ib.ch</a> ).
CC		-----
DR	EMBL:	AF067864; A045561.1; -
DR	EMBL:	AF053356; AAC8796.1; -
DR	EMBL:	AK022002; BAB13951.1; -
DR	EMBL:	AK000421; BAA91153.1; ALT_INIT.
DR	MEROPS:	M28_973; -
DR	MIM:	604720; -
DR	MIM:	604250; -
DR	InterPro:	IPIR003137; PA.
DR	Pfam:	PF02225; PA; 1.
KW	Transmembrane;	Glycoprotein; Receptor; Signal-anchor;
KW	Alternative splicing;	Disease mutation.
KW	DOMAIN	1 83
FT	TRANSMEM	84 104
FT	FT	
FT	FT	
FT	DOMAIN	105 801
FT	SITE	23 26
FT	DISULFID	108 108
FT	DISULFID	111 111
FT	CARBOHYD	240 240
FT	CARBOHYD	339 339
FT	CARBOHYD	540 540
FT	CARBOHYD	754 754
FT	VARSPLIC	1 171
FT	VARSPLIC	343 369
FT	VARIANT	172 172
FT	FT	
FT	CONFLICT	712 712
SO	SEQUENCE	801 AA; 88755 MM; DDD3082BA835413A CRC64;
Query Match 68.0%; Score 34; DB 1; Length 801;		
Best Local Similarity 75.0%; Pred. No. 31;		
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;		
Oy	2 RGIRPYGR 9	
	:	
Db	262 RGVDPVGR 269	
RESULT 10		
PPOX_HUMAN		
ID PPOX_HUMAN	STANDARD;	PRT; 477 AA.
AC P50336;		
DT 01-OCT-1996 (Rel. 34, Created)		
DT 01-OCT-1996 (Rel. 34, Last sequence update)		
DT 15-DEC-1998 (Rel. 37, Last annotation update)		
DE Protoporphyryrrogen oxidase (EC 1.3.3.4) (PRO).		
GN PPOX.		
OS Homo sapiens (Human).		
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.		
NCBI_TaxID=9606;		
NN [1]		

RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE-95229621; PubMed-7713909;  
 RA Nishimura K., Taketani S., Inokuchi H.;  
 RT "Cloning of a human cDNA for protoporphyrinogen oxidase by  
 RT complementation in vivo of a hemg mutant of Escherichia coli.";  
 RL J. Biol. Chem. 270:8076-8080(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE-96367087; PubMed-8771201;  
 RA Dailey T.A., Dailey H.A.;  
 RT "Human protoporphyrinogen oxidase: expression, purification, and  
 RT characterization of the cloned enzyme.";  
 RL Protein Sci. 5:98-105(1996).  
 [3]  
 RP VARIANT VP ARG-232, AND VARIANT HIS-304.  
 RX MEDLINE-97005368; PubMed-8852667;  
 RA Deybach J.-C., Puy H., Robreau A.-M., Lamoril J., da Silva V.,  
 RA Grandchamp B., Nordmann Y.;  
 RT "Mutations in the protoporphyrinogen oxidase gene in patients with  
 RT variegate porphyria.";  
 RL Hum. Mol. Genet. 5:407-410(1996).  
 [4]  
 RP VARIANT VP TRP-59, AND VARIANT CYS-168.  
 RX MEDLINE-96241580; PubMed-8673113;  
 RA Weisner P.N., Dailey T.A., Hift R.J., Ziman M., Corrigan A.V.,  
 RA Roberts A.G., Weisner D.M., Kirsch R.E., Dailey H.A.;  
 RT "A R59W mutation in human protoporphyrinogen oxidase results in  
 RT decreased enzyme activity and is prevalent in South Africans with  
 RT variegate porphyria.";  
 RL Nat. Genet. 13:95-97(1996).  
 [5]  
 RP VARIANT VP CYS-152.  
 RX MEDLINE-98434271; PubMed-9763307;  
 RA Frank J., Poh-Fitzpatrick M.B., King L.E. Jr., Christian A.M.;  
 RT "The genetic basis of 'Scarsdale Gourmet Diet' variegate porphyria: a  
 RT missense mutation in the protoporphyrinogen oxidase gene.";  
 RL Arch. Dermatol. Res. 290:441-445(1998).  
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN  
 CC IX TO FORM PROTOPORPHYRIN IX.  
 CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) = protoporphyrin-  
 CC IX + H(2)O(2).  
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.  
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE  
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,  
 CC LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.  
 CC -1- DISEASE: DEFECTS IN POX ARE THE CAUSE OF PORPHYRIA VARIEGATA  
 CC (VP), A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR  
 CC HYPERTRICHOSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE  
 CC INTERMITTENT PORPHYRIA.  
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 CC DR EMBL: D38537; BAA07538.1; -;  
 DR EMBL: U26446; AAA67690.1; -;  
 DR MIM: 600923; -;  
 DR MIM: 176200; -;  
 KW Porphyrin biosynthesis; Heme biosynthesis; Oxidoreductase;  
 KW Flavoprotein; FAD; Mitochondrion; Disease mutation; Polymorphism.  
 FT NP\_BIND 9 14 FAD (POTENTIAL).  
 FT VARIANT 59 59 R -> W (IN VP).  
 FT /ftid-var\_003686.

FT VARIANT 152 152 R -> C (IN VP).  
 FT /ftid-var\_003687.  
 FT VARIANT 168 168 R -> C.  
 FT /ftid-var\_003688.  
 FT VARIANT 232 232 G -> R (IN VP).  
 FT /ftid-var\_003689.  
 FT VARIANT 304 304 R -> H.  
 FT /ftid-var\_003690.  
 SQ SEQUENCE 477 AA; 50765 MW; 2444DEAC2E6C33EE CRC64;  
 Query Match 66.0%; Score 33; DB 1; Length 477;  
 Best Local Similarity 85.7%; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 2 RGRPYG 8  
 Db 59 RGRPYG 65  
 RESULT 11  
 PROX\_MOUSE  
 ID PROX\_MOUSE STANDARD; PRT; 477 AA.  
 AC P51175; P97344;  
 DT 01-OCT-1996 (rel. 34, Created)  
 DT 01-OCT-1996 (rel. 34, Last sequence update)  
 DT 15-DEC-1998 (rel. 37, Last annotation update)  
 DE Protoporphyrinogen oxidase (EC 1.3.3.4) (PPO).  
 GN PROX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96132666; PubMed-8554330;  
 RA Dailey T.A., Dailey H.A., Weisner P., Prasad A.R.;  
 RT "Cloning, sequence, and expression of mouse protoporphyrinogen  
 RT oxidase.";  
 RL Arch. Biochem. Biophys. 324:379-384(1995).  
 [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Erythrocyte;  
 RX MEDLINE-95331315; PubMed-7607249;  
 RA Taketani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,  
 RA Nishimura K., Inokuchi H.;  
 RT "Induction of terminal enzymes for heme biosynthesis during  
 RT differentiation of mouse erythroleukemia cells.";  
 RL Eur. J. Biochem. 230:760-765(1995).  
 [3]  
 RP SUBCELLULAR LOCATION.  
 RC TISSUE-Liver;  
 RX MEDLINE-88153682; PubMed-3346226;  
 RA Ferreira G.C., Andrew T.L., Karr S.W., Dailey H.A.;  
 RT "Organization of the terminal two enzymes of the heme biosynthetic  
 RT pathway. Orientation of protoporphyrinogen oxidase and evidence for a  
 RT membrane complex.";  
 RL J. Biol. Chem. 263:3835-3839(1988).  
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN  
 CC IX TO FORM PROTOPORPHYRIN IX.  
 CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) = protoporphyrin-  
 CC IX + H(2)O(2).  
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).  
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE  
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.  
 CC -1- INDUCTION: DURING ERYTHROID DIFFERENTIATION.  
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U25114; AAA96003.1; -  
DR EMBL: D45185; BAA08126.1; -  
DR MGD: MGI:104968; Ppx.  
KW Porphyria biosynthesis; Heme biosynthesis; Oxidoreductase;  
KM Flavoprotein; FAD; Mitochondrion.  
FT NP\_BIND 9 14 FAD (POTENTIAL).  
FT CONFLICT 64 64 A -> T (IN REF. 2).  
FT CONFLICT 66 66 A -> P (IN REF. 2).  
FT CONFLICT 108 108 L -> S (IN REF. 2).  
FT CONFLICT 427 427 W -> C (IN REF. 2).  
SQ SEQUENCE 477 AA; 50870 MW; 8CFB48120728DE6F CRC64;

Query Match 66.0%; Score 33; DB 1; Length 477;  
Best Local Similarity 85.7%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RGIRPVG 8  
Db 59 RGIRPAG 65

## RESULT 12

SYE\_BACHD STRAND: PRT; 485 AA.

AC Q9KGF6; 16-OCT-2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Glutamy-tRNA synthetase (EC 6.1.1.17) (glutamate--tRNA ligase)

DE (GLURS).

GN GLTX OR BH0109.

OS Bacillus halodurans.

OC Bacteria; Firmicutes; Bacillus/Clostridium group;

OC Bacillus/Staphylococcus group; Bacillus.

OX NCBI\_TaxID=86665;

RN [1]

RP SEQUENCE FROM N.A.

RC STRAIN=C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132;

RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,

RA Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,

RA Horikoshi K.;

RT "Complete genome sequence of the alkaliphilic bacterium Bacillus

RT halodurans and genomic sequence comparison with Bacillus subtilis.";

RL Nucleic Acids Res. 28:4317-4331(2000).

CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP +

CC diphosphate + L-glutamyl-tRNA(Glu).

CC -1- SUBUNIT: MONOMER (BY SIMILARITY).

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: AP001507; BAB03828.1; -

DR InterPro: IPR000924; tRNA-synt\_1c.

DR InterPro: IPR001412; tRNA-synt\_1c.

DR Pfam: PF00749; tRNA-synt\_1c; 1.

DR PRINTS: PR00987; TRNASYNTHGLU.

DR PROSITE: PS00178; AA\_TRNA\_LIGASE\_I. 1.

KW Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;

KW Complete proteome.

FT SITE 11 21 "HIGH" REGION.  
FT SITE 252 256 "RMSKS" REGION.  
FT BINDING 255 255 ATP (BY SIMILARITY).  
SQ SEQUENCE 485 AA; 54785 MW; 7D34A862918F5B6 CRC64;

Query Match 66.0%; Score 33; DB 1; Length 485;  
Best Local Similarity 66.7%; Pred. No. 30;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
Db 149 KGIRPVGRF 157

## RESULT 13

ULN7\_HCMVA STRAND: PRT; 150 AA.

AC P16828; 01-AUG-1990 (Rel. 15, Created)

DT 01-AUG-1990 (Rel. 15, Last sequence update)

DT 01-FEB-1991 (Rel. 17, Last annotation update)

DE Hypothetical protein ULN07.

GN ULN07.

OS Human cytomegalovirus (strain AD169).

OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;

OC Betaherpesvirinae; Cytomegalovirus.

OX NCBI\_TaxID=10360;

RN [1]

RP SEQUENCE FROM N.A.

RX MEDLINE=90269039; PubMed=2161319;

RA Chee M.S., Bankier A.T., Beck S., Bohnl R., Brown C.M., Cerny R.,

RA Horsnell T., Hutchinson C.A., Ili, Kouzarides T., Martignetti J.A.,

RA Preddie E., Satchell S.C., Tomlinson P., Weston K.M., Barrell B.G.;

RT "Analysis of the protein-coding content of the sequence of human

RT cytomegalovirus strain AD169.";

RT Curr. Top. Microbiol. Immunol. 154:125-169(1990).

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC EMBL: X17403; CAA35343.1; -

DR PIR: S09872; S09872.

KW Hypothetical protein.

FT CARBOHYD 144 144 N-LINKED (GLCNAC... ) (POTENTIAL).

SQ SEQUENCE 150 AA; 17373 MW; 042707546C7EB878 CRC64;

Query Match 64.0%; Score 32; DB 1; Length 150;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 2 RGIRPVGRF 10  
Db 57 RGQRPVGRF 65

## RESULT 14

Y361\_SYNY3 STRAND: PRT; 249 AA.

AC Q55578; 15-DEC-1998 (Rel. 37, Created)

DT 15-DEC-1998 (Rel. 37, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Hypothetical 28.2 kDa protein slr0361.

GN SLR0361.

OS Synechocystis sp. (strain PCC 6803).

OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

```
OX NCB1_TaxID=1148;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166(1995).
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF PSEUDOURIDINE SYNTHASES.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; D63999; BAA10082.1; -
CC InterPro: IPR000613; Pseudou_synth.
CC InterPro: IPR000748; Pal_RSU.
CC InterPro: IPR002942; S4.
CC Pfam; PF00849; Pseudou_synth_2; 1.
CC Pfam; PF04479; S4; 1.
CC SMART; SM00363; S4; 1.
CC PROSITE; PS01149; Psi_RSU; 1.
CC Hypothetical protein; Complete proteome.
KW SEQUENCE 249 AA; 28228 MW; 315F208030180326 CRC64;
SO

Query Match 64.0%; Score 32; DB 1; Length 249;
Best Local Similarity 62.5%; Pred. No. 24;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GIRPYGR 9
DB 100 KGLHPVGR 107

RESULT 15
AR72_HUMAN
ID AR72_HUMAN STANDARD; PRT; 330 AA.
AC 043488; 075749;
DT 15-JUL-1999 (Rel. 38, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aflatoxin B1 aldehyde reductase 1 (EC 1.-.-.-) (AFBI-AR 1)
DE (Aldoketoreductase 7).
GN AKR7A2 OR AFAR OR AKR7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCB1_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RP TISSUE=Liver;
RC MEDLINE=98244807; PubMed=9576847;
RA Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;
RT "Molecular cloning, expression and catalytic activity of a human AKR7
RT major of the aldo-keto reductase superfamily: evidence that the
RT major 2-carboxybenzaldehyde reductase from human liver is a homologue
RT of rat aflatoxin B1-aldehyde reductase.";
RL Biochem. J. 332:21-34(1998).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=99040634; PubMed=9823300;
RA Pramli C., Savelyeva L., Perri P., Schwab M.;
RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-
RT 1p36.1 in a region frequently altered in human tumor cells.";
RL Cancer Res. 58:5014-5018(1998).
```

```
RN [3]
RP SEQUENCE FROM N.A.
RA Hall R.;
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOLCHOL. COULD BE
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC
CC EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE
CC REDUCTASE.
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; AF026947; AAC52104.1; -
CC EMBL; Y16675; CAA76347.1; -
CC EMBL; AL035413; CAB72321.1; -
CC MIM: 603418; -
CC InterPro: IPR001395; Aldo_ket_red.
CC Pfam; PF00248; aldo_ket_red; 1.
CC Oxidoreductase.
FT ACT SITE 112 112 HYDROGEN-BOND DONOR (PROBABLE).
FT CONFLICT 113 113 A -> T (IN REF. 1).
FT SEQUENCE 330 AA; 36618 MW; 3BBFB7EDDCARF4D54 CRC64;
SO
```

```
Query Match 64.0%; Score 32; DB 1; Length 330;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
```

```
OY 3 GIRPYGR 10
DB 215 GKQPVGRF 222
```

Search completed: September 13, 2002, 09:30:48  
Job time: 1138 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

## OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:23 : Search time 311.85 Seconds  
(without alignments)  
6.102 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_11\_21  
Perfect score: 50  
Sequence: 1 XRGIRPVGRFX 11

Scoring table:  
BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 segs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

1: SPREMBL\_19:\*  
2: sp\_archaea:\*  
3: sp\_bacteria:\*  
4: sp\_fungi:\*  
5: sp\_human:\*  
6: sp\_invertebrate:\*  
7: sp\_mammal:\*  
8: sp\_mhc:\*  
9: sp\_organelle:\*  
10: sp\_phage:\*  
11: sp\_plant:\*  
12: sp\_rodent:\*  
13: sp\_virus:\*  
14: sp Vertebrate:\*  
15: sp\_uniclassified:\*  
16: sp\_rvrius:\*  
17: sp\_bacteriap:\*  
17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	46	92.0	117	13	Q9W624	Q9W624 carassius a
2	43	86.0	420	2	Q931Z7	Q931Z7 streptomyce
3	36	72.0	426	17	Q979S0	Q979S0 thetoplasma
4	36	72.0	674	16	Q50431	Q50431 mycobacteri
5	35	70.0	272	3	O42924	O42924 schizosacch
6	35	70.0	371	16	Q981V3	Q981V3 rhizobium l
7	35	70.0	504	16	O53895	O53895 mycobacteri
8	35	70.0	519	16	O925G7	O925G7 mycobacteri
9	35	70.0	894	16	Q9KSJ8	Q9KSJ8 vibrio chol
10	35	70.0	1000	6	Q9V840	Q9V840 drosophila
11	35	70.0	1056	16	Q9PAQ0	Q9PAQ0 xylella fas
12	35	70.0	1172	5	Q9VU79	Q9VU79 drosophila
13	34	68.0	329	2	Q9RJ77	Q9RJ77 streptomyce
14	34	68.0	529	3	Q9C1M8	Q9C1M8 pleurotus s
15	34	68.0	537	4	O75230	O75230 homo sapien
16	34	68.0	780	4	O73422	O73422 homo sapien

17	34	68.0	801	4	Q9UP52	Q9UP52 homo sapien
18	33	66.0	139	5	Q9N6S1	Q9N6S1 leishmania
19	33	66.0	153	9	Q9T133	Q9T133 bacterioph
20	33	66.0	205	4	Q96RC9	Q96RC9 homo sapien
21	33	66.0	253	2	P72459	P72459 streptomyce
22	33	66.0	313	16	Q92Y28	Q92Y28 rhizobium m
23	33	66.0	319	2	O88067	O88067 streptomyce
24	33	66.0	335	16	P96221	P96221 mycobacteri
25	33	66.0	353	16	O33194	O33194 mycobacteri
26	33	66.0	366	16	O66950	O66950 aquifex aeo
27	33	66.0	445	2	O50531	O50531 streptomyce
28	33	66.0	477	11	Q99M34	Q99M34 mus musculu
29	33	66.0	511	3	O59714	O59714 schizosacch
30	33	66.0	589	16	Q9A718	Q9A718 caulobacter
31	33	66.0	883	10	Q9LEG6	Q9LEG6 lycopersico
32	33	66.0	901	16	Q99YPS	Q99YPS streptococc
33	33	66.0	1940	16	Q9RUC8	Q9RUC8 deinococcus
34	33	66.0	2104	5	Q21281	Q21281 caenorhabdl
35	33	66.0	2104	5	Q964N4	Q964N4 caenorhabdl
36	33	66.0	2104	8	Q9B5K5	Q9B5K5 balanus gla
37	32	64.0	154	2	O54139	O54139 streptomyce
38	32	64.0	183	16	Q92KK8	Q92KK8 rhizobium m
39	32	64.0	183	17	Q979Z6	Q979Z6 thermoplasma
40	32	64.0	266	8	Q23888	Q23888 dictyostell
41	32	64.0	278	16	Q92QR1	Q92QR1 rhizobium m
42	32	64.0	285	2	Q9X4W7	Q9X4W7 pseudomonas
43	32	64.0	292	16	Q92VM1	Q92VM1 rhizobium m
44	32	64.0	299	16	P95060	P95060 mycobacteri
45	32	64.0	309	2	O86546	O86546 streptomyce

## ALIGNMENTS

RESULT 1  
ID Q9W624 PRELIMINARY; PRT; 117 AA.  
AC Q9W624;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE C-RF AMIDE.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Satake H., Minakata H., Fujimoto M.;  
RT "Carassius Rhamde (C-RF amide)".  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB020024; BAA76662.1; -  
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 92.0%; Score 46; DB 13; Length 117;  
Best Local Similarity 77.8%; Pred. No. 0.11;  
Matches 7; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 2 RGRPVGRF 10  
11:11:111  
Db 67 RGRPVGRF 75  
RESULT 2  
ID Q931Z7 PRELIMINARY; PRT; 420 AA.  
AC Q931Z7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
CC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-CM3239;  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CM3239."  
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AF033994; AAK61719.1; -  
SO SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;

Query Match 86.0%; Score 43; DB 2; Length 420;  
Best Local Similarity 100.0%; Pred. NO. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 GIRPYGRF 10  
DB 48 GIRPYGRF 55  
|||||

RESULT 3  
O979S0 PRELIMINARY; PRT; 426 AA.  
ID O979S0;  
AC O979S0;  
DT 01-OCT-2001 (TREMBLrel. 18. Created)  
DT 01-OCT-2001 (TREMBLrel. 18. Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19. Last annotation update)  
DE TVG1120468 PROTEIN.  
GN TVG1120468.  
OS Thermoplasma volcanium.  
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmataceae;  
CC Thermoplasma.  
OX NCBI\_TaxID=50339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-GSSI / DSM 4299 / JCM 9571;  
RX MEDLINE=20570466; PubMed=11121031;  
RA Kawashima T., Ameno N., Koike H., Makino S.-I., Higuchi S.,  
RA Kawashima-Ohya Y., Watanabe K., Yamazaki K., Kaneshiro K., Kawamoto T.,  
RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
RT "Archaeal adaptation to higher temperatures revealed by genomic  
RT sequence of Thermoplasma volcanium."  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
DR EMBL: AP000995; BAB60232.1; -  
DR InterPro: IPR002881; DUF58.  
DR InterPro: IPR002035; VMFA.  
DR Pfam: PF01882; DUF58; 1.  
DR SMART: SMO0327; VMA; 1.  
KW Complete proteome.  
SO SEQUENCE 426 AA; 49092 MW; 63967CF711116F4A CRC64;

Query Match 72.0%; Score 36; DB 17; Length 426;  
Best Local Similarity 75.0%; Pred. NO. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 GIRPYGRF 10  
DB 304 GIRPYGRF 311  
|||||

RESULT 4  
O50431 PRELIMINARY; PRT; 674 AA.  
ID O50431;  
AC O50431;  
DT 01-JUN-1998 (TREMBLrel. 06. Created)  
DT 01-JUN-1998 (TREMBLrel. 06. Last sequence update)

DT 01-DEC-2001 (TREMBLrel. 19. Last annotation update)  
DE 2,4-DIENYL-COA REDUCTASE.  
GN FADH OR RV1175C OR MTW005.11C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
CC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Fellwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., Mclean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
RT complete genome sequence."  
RL Nature 393:537-544(1998).  
DR EMBL: AL010166; CAAL5852.1; -  
DR Tuberculist; RV1175C; -  
DR InterPro: IPR001064; Crystallin.  
DR InterPro: IPR001327; FAD\_pyr\_redox.  
DR InterPro: IPR000205; NAD\_binding.  
DR InterPro: IPR001155; Oxidored\_FMN.  
DR Pfam: PF00724; oxidored\_FMN; 1.  
DR Pfam: PF00070; pyr\_redox; 1.  
DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
KW Complete proteome.  
SO SEQUENCE 674 AA; 72901 MW; 8F478F74ED3E0A5 CRC64;

Query Match 72.0%; Score 36; DB 16; Length 674;  
Best Local Similarity 85.7%; Pred. NO. 75;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 GIRPYGR 9  
DB 492 GIRPYGR 498  
|||||

RESULT 5  
O42924 PRELIMINARY; PRT; 272 AA.  
ID O42924;  
AC O42924;  
DT 01-JAN-1999 (TREMBLrel. 09. Created)  
DT 01-JAN-1999 (TREMBLrel. 09. Last sequence update)  
DT 01-JUN-2001 (TREMBLrel. 17. Last annotation update)  
DE HYPOTHETICAL 30.9 KDA PROTEIN C16A3.19 IN CHROMOSOME II.  
GN SPG016A3.19.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
CC Schizosaccharomycetes.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL021748; CAAL6870.1; -  
DR InterPro: IPR001005; MYB\_DNA\_bind.  
DR PROSITE: PS00037; MYB\_1; UNKNOWN\_1.  
KW Hypothetical protein.  
SO SEQUENCE 272 AA; 30901 MW; 8FE0EB29217107A8 CRC64;

Query Match 70.0%; Score 35; DB 3; Length 272;  
Best Local Similarity 85.7%; Pred. NO. 45;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

Oy      2  RGIRPVGR 8
      11:11111
Db      47  RGLRPVG 53

RESULT 6
O981V3 PRELIMINARY; PRT; 371 AA.
AC 0981V3:
DT 01-OCT-2001 (TrEMBLrel. 18, Created)
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DE MUR9222 PROTEIN.
GN MUR9222.
OS Rhizobium loti (Mesorhizobium loti).
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF30309;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shimpō S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003015; BAB54606.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 371 AA; 42385 MW; 98D9874B5E7831D1 CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 16; Length 371;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Oy      2  RGIRPVGR 9
      11:11111
Db      179  RGEVPIGR 186

RESULT 7
O53895 PRELIMINARY; PRT; 504 AA.
AC 053895:
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE TWO-COMPONENT SENSOR.
GN RV0982 OR MTV044.10.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC MEDLINE=98295987; PubMed=9634230;
RX Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eigjmeier K., Gas S., Barry C.E. III, Tekala F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holtroyd S.,
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Rutter S., Seeger K., Skelton S., Squares R.,
RA Sultson J.E., Taylor K., Whitehead S., Barrell B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."

```

```

RL      Nature 393:537-544(1998).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL: AL021999; CAI7581.1; -.
DR HSSP: P02933; 1BXD.
DR Tuberculin; RV0982; -.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_KinA.
DR InterPro: IPR004359; His_Kin_519.
DR Pfam: PF00672; HAMF; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; signal; 1.
DR SMART: SM00304; HAMF; 1.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00389; HisK_A; 1.
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;
KW Transference.
SQ SEQUENCE 504 AA; 54407 MW; 26792A71AC432232 CRC64;

Query Match
Best Local Similarity 70.0%; Score 35; DB 16; Length 504;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Oy      3  GIRPVGR 9
      11:11111
Db      188  GLRPVGR 194

RESULT 8
O925G7 PRELIMINARY; PRT; 519 AA.
ID 0925G7:
AC 0925G7:
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)
DE PUTATIVE TWO-COMPONENT SYSTEM SENSOR KINASE.
GN MLO175 OR MLCB373.27.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris D., Taylor K.;
RL Submitted (Feb-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eigjmeier K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Duthoy S., Feltwell T., Fraser A., Hamlin N.,
RA Holtroyd S., Hornsby T., Jagels K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Rutter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrell B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL: AL035500; CAB36689.1; -.
DR EMBL: AL583917; CAC29683.1; -.
DR HSSP: P02933; 1BXD.
DR Leproma; MLO175; -.
DR InterPro: IPR004358; BCTRLSENSOR.
DR InterPro: IPR003660; HAMF.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_KinA.
DR InterPro: IPR004359; His_Kin_519.
DR Pfam: PF00672; HAMF; 1.

```

DR Pfam: PF00512; signal; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.  
DR SMART: SM00304; HAMF; 1.  
DR SMART: SM00387; HATPase\_C; 1.  
DR SMART: SM00388; HSKA; 1.  
KW Complete proteome; kinase; phosphorylation; sensory transduction;  
KW Transferrase  
SO SEQUENCE 519 AA; 56470 MW; C0F2DA4D57364F10 CRC64;

Query Match 70.0%; Score 35; DB 16; Length 519;  
Best Local Similarity 85.7%; Pred. No. 90;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 GRIPVGR 9  
||:|||||  
Db 188 GRIPVGR 194

RESULT 9  
O9KSJ8 PRELIMINARY; PRT; 894 AA.  
AC O9KSJ8;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
DE DNA GYRASE, SUBUNIT A.  
GN VC1258.  
OS *Vibrio cholerae*.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; *Vibrio*.  
OX NCBI\_TaxID=666;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR M16961 / SEROTYPE O1;  
RX MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ervolaeva M.D., Vamathevan J., Bass S., Qin H., Dragol I., Sellers P.,  
RA McDonald L., Ueberback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen *Vibrio*  
RT *cholerae*.";  
RL Nature 406:477-483(2000).  
DR EMBL: AE004205; AAF94417.1; -.  
DR HSSP: P09097; IAB4.  
DR TIGR: VC1258; -.  
DR InterPro: IPR002205; DNA\_topoisomIV.  
DR Pfam: PF00521; DNA\_topoisomIV; 1.  
DR SMART: SM00434; TOP4C; 1.  
KW Complete proteome.  
SO SEQUENCE 894 AA; 99025 MW; C7EA42056A08F57F CRC64;

Query Match 70.0%; Score 35; DB 16; Length 894;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 2 GRIPVGR 9  
||:|||||  
Db 733 GRIPVGR 740

RESULT 10  
O9VS40 PRELIMINARY; PRT; 1000 AA.  
AC O9VS40;  
DT 01-MAY-2000 (TREMblrel. 13, Created)  
DT 01-MAY-2000 (TREMblrel. 13, last sequence update)  
DT 01-DEC-2001 (TREMblrel. 19, last annotation update)  
DE CG8608 PROTEIN.  
GN UNC-13-4A OR CG8608.

OS *Drosophila melanogaster* (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; *Drosophilidae*; *Drosophila*.  
OX NCBI\_TaxID=7227;  
[1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amentides P.G., Scherer S.E., Li P.W., Hoiskins R.A., Galie R.F.,  
RA Amentides P.G., Scherer S.E., Li P.W., Hoiskins R.A., Galie R.F.,  
RA Sutton G.G., Wortman J.R., Richards S., Ashburner M., Henderson S.N.,  
RA Brandon R.C., Rogers Y.-H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Abrell J.F., Agbayani A., An H.-J., Andrews-Pfankoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brockstein P., Brotler P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glodet A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jalelli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Klamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveli J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of *Drosophila melanogaster*.";  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003559; AAF50589.1; -.  
DR HSSP: P21707; IBYN  
DR FLYbase: FBgn0035756; unc-13-4A.  
DR InterPro: IPR000008; C2.  
DR Pfam: PF00168; C2; 1.  
DR PRINTS: PR00360; C2DOMAIN.  
DR SMART: SM00239; C2; 2.  
DR PROSITE: PSS0004; C2\_DOMAIN\_2; 1.  
SO SEQUENCE 1000 AA; 114286 MW; FDF14F05A55696A2 CRC64;

Query Match 70.0%; Score 35; DB 5; Length 1000;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 2 GRIPVGR 10  
||:|||||  
Db 130 GRIPVGR 138

RESULT 11  
O9PAQ0 PRELIMINARY; PRT; 1056 AA.  
AC O9PAQ0;  
DT 01-OCT-2000 (TREMblrel. 15, Created)  
DT 01-OCT-2000 (TREMblrel. 15, last sequence update)



DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
GN HYPOTHEICAL PROTEIN XF2445.  
DE XF2445.  
OS Xylella fastidiosa.  
OC Bacteria: Proteobacteria: gamma subdivision: Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=9A5C;  
RX MEDLINE=20365717; PubMed=10910347;  
RA Simpson A.J.G., Reimach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvarenga R., Alves L.M.C., Araya J.E., Bala J.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carier H.,  
RA Colatto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facinanci A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Froime M., Furian L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohnsels J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambis M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Lemos M.A., Madeira A.M.B.N., Madeira H.M.F., Matsukuma A.Y.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesqueto J.B.,  
RA Quaggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsubako M.H.,  
RA Varella H., Van Sluys M.A., Vertovskii Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT The genome sequence of the plant pathogen Xylella fastidiosa.;  
RL Nature 406:151-159(2000).  
DR EMBL: AE004053; AAF85244.1; -  
DR InterPro: IPR000421; FAS58.C.  
DR InterPro: IPR001092; HLH\_dlm.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 1056 AA; 118471 MW; 361CF2510BDFEC10 CRC64;

Query Match 70.0%; Score 35; DB 16; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 3 GINPVGR 9  
Db 1047 GLRPVGR 1053  
1:|||||

RESULT 12  
O9VU79 PRELIMINARY; PRT; 1172 AA.  
AC O9VU79;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE CG10738 PROTEIN.  
GN CG10738.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;

RA Adams M.D., Celisner S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Mortan J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers J.H.C., Blazey R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miles G.L.G.,  
RA April J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brothier P.,  
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Dou P.L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin R.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Fostler C., Gabrielian A.E., Gang N.S., Gelbart W.M., Glasser K.,  
RA Glodok A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
RA Jostali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum R.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Mishina N.V., Mobarry C., Morris J., Mostrel A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Svirskas R., Tector C., Turner R., Venter L., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasserman D.A., Weinstein G.M., Welschbach J.,  
RA Williams S.M., Woodard T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT The genome sequence of Drosophila melanogaster.;  
RL Science 287:2185-2195(2000).  
DR EMBL: AE003537; AAF49811.1; -  
DR HSSP: 002846; 1AHL.  
DR rYBase: FBgn0036368; CG10738.  
DR InterPro: IPR001828; ANF\_receptor.  
DR InterPro: IPR000719; Euk\_kinase.  
DR InterPro: IPR001054; Guanylate\_cyclase.  
DR Pfam: PF01094; ANF\_receptor; 2.  
DR Pfam: PF00211; guanylate\_cyc; 1.  
DR Pfam: PF00069; kinase; 1.  
DR SMART: SM00044; CYC; 1.  
DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
DR PROSITE: PS50125; GUANYLATE\_CYCLASES\_2; 1.  
DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
DR ATP-binding; lyase; transferase.  
SQ SEQUENCE 1172 AA; 131992 MW; C6450931F82A6D55 CRC64;

Query Match 70.0%; Score 35; DB 5; Length 1172;  
Best Local Similarity 75.0%; Pred. No. 2.2e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 3 GINPVGR 10  
Db 434 GLRPVGR 441  
1:|||||

RESULT 13  
O9RUR7 PRELIMINARY; PRT; 329 AA.  
AC O9RUR7;  
DT 01-MAY-2000 (TrEMBLrel. 13, Created)  
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE PUTATIVE ZINC-BINDING OXIDOREDUCTASE.  
GN SCF51.18.

OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomyces; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Seeger K.V., Harris D.;  
 RL Submitted (OCT-1999) to the EMBL/Genbank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
 RN 13  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA MEDLINE=97000351; PubMed=8843436;  
 RA Redenbach M., Kleser H.M., Denaplatte D., Elchner A., Cullum J.,  
 RA Kinashl H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL MOL. Microbiol. 21:77-96(1996).  
 DR EMBL: AL132707; CAB59716.1; -  
 DR InterPro: IPR002085; Adh\_zn\_family.  
 DR Pfam: PF00107; adh\_zinc.1.  
 SQ SEQUENCE 329 AA; 34155 MM; 5DCAD4FB174FD042 CRC64;

Query Match 68.0%; Score 34; DB 2; Length 329;  
 Best Local Similarity 87.5%; Pred. No. 88;  
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RGIRPYGR 9  
 DB 137 RGIRPYGR 144

RESULT 14  
 O9C1M8 PRELIMINARY; PRT; 529 AA.  
 AC O9C1M8;  
 DT 01-JUN-2001 (TREMBLrel. 17, Created)  
 DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 DE HEAT-INDUCED CATALASE.  
 OS Pleurotus sajor-caju (Oyster mushroom).  
 OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
 OC Agaricales; Pleurotaceae; Pleurotus.  
 OX NCBI\_TaxID=50053;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Jeong M.-J., Park S.-C.;  
 RT "Cloning of the catalase gene from Pleurotus sajor-caju";  
 RL Submitted (JUL-2000) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AF286097; AK15159.1; -  
 DR HSPF; P15202; 1A4E.  
 DR InterPro: IPR002226; Catalase.  
 DR Pfam: PF00199; catalase; 1.  
 DR PRINTS: PR00067; CATALASE.  
 DR ProDom: PD000510; Catalase; 1.  
 DR PROSITE: PS00437; CATALASE\_1; 1.  
 DR PROSITE: PS00438; CATALASE\_2; UNKNOWN\_1.  
 SQ SEQUENCE 529 AA; 59791 MM; A432B6445B6AC13 CRC64;

Query Match 68.0%; Score 34; DB 3; Length 529;  
 Best Local Similarity 71.4%; Pred. No. 1.5e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 4 IRPVGRF 10  
 DB 298 LRPIGRF 304

RESULT 15  
 ID 075230 PRELIMINARY; PRT; 537 AA.  
 AC 075230;  
 DT 01-NOV-1998 (TREMBLrel. 08, Created)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last sequence update)  
 DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
 DE H53\_GSI (FRAGMENT).  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN 11  
 RP SEQUENCE FROM N.A.  
 RA Kimmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,  
 RA Kader K., Miguel T., Miller C., Piluck S., Pollard M., Rojeski H.,  
 RA Subramanian S., Martin C.H.;  
 RT "Sequencing of human chromosome 5";  
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.  
 RN 12  
 RP SEQUENCE FROM N.A.  
 RA Rieke D.O.;  
 RT "Large Scale Sequence Analysis and Annotation with the Sequence  
 RT Comparison Analysis (SCAN) System";  
 RL Submitted (JUN-1998) to the EMBL/Genbank/DBJ databases.  
 DR EMBL: AC005178; AAC24489.1; -  
 FT NON\_TER 1  
 FT 537  
 SQ SEQUENCE 537 AA; 58017 MM; 2ED384E9DB64BC5 CRC64;

Query Match 68.0%; Score 34; DB 4; Length 537;  
 Best Local Similarity 75.0%; Pred. No. 1.5e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 RGIRPYGR 9  
 DB 295 RGIRPYGR 302

Search completed: September 13, 2002, 09:29:24  
 Job time: 1069 sec.

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

# OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:36 ; Search time 399.68 Seconds  
(without alignments)  
2.779 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_12\_21

Perfect score: 49  
Sequence: 1 RGIKPVGRFX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

## Database :

A.Geneseq.032802.\*  
1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	98.0	10	20	AAW95178
2	48	98.0	15	18	AAW31400
3	48	98.0	15	20	AAW97230
4	48	98.0	15	21	AAV49293
5	48	98.0	15	21	AAV49295
6	48	98.0	15	21	AAV49296
7	48	98.0	20	18	AAW31394
8	48	98.0	20	18	AAW31387
9	48	98.0	20	18	AAW31374
10	48	98.0	20	18	AAW97232
11	48	98.0	20	20	AAW97234

12	48	98.0	20	20	AAW97236	Human type ligand
13	48	98.0	20	20	AAW95191	Bovine pituitary-d
14	48	98.0	20	20	AAW95175	Murine pituitary-d
15	48	98.0	20	21	AAW10350	Bovine oxytocin se
16	48	98.0	20	21	AAW10358	Rat oxytocin secre
17	48	98.0	20	21	AAW10365	Human oxytocin sec
18	48	98.0	20	21	AAW10369	Oxytocin secretion
19	48	98.0	20	21	AAW49294	1992 ligand peptid
20	48	98.0	20	21	AAW49301	1992 ligand peptid
21	48	98.0	20	21	AAW49302	1992 ligand peptid
22	48	98.0	20	22	AAW62519	Bovine CRH releas
23	48	98.0	20	22	AAW62527	Rat CRH releas
24	48	98.0	20	22	AAW62534	Human CRH releas
25	48	98.0	20	22	AAW62538	CRH releasing prot
26	48	98.0	20	22	AAW90992	Proactin releasin
27	48	98.0	20	22	AAW90994	Proactin releasin
28	48	98.0	20	22	AAW69596	Peptide PRP20 fra
29	48	98.0	20	22	AAW69594	Synthetic G protei
30	48	98.0	21	18	AAW31397	Human type G prote
31	48	98.0	21	18	AAW31395	Human type G prote
32	48	98.0	21	18	AAW31388	Rat type G protein
33	48	98.0	21	18	AAW31375	Bovine G protein-c
34	48	98.0	21	20	AAW97227	Partial ligand pol
35	48	98.0	21	20	AAW87616	Mammalian 1992 lig
36	48	98.0	21	20	AAW95192	Bovine pituitary-d
37	48	98.0	21	21	AAW10351	Bovine oxytocin se
38	48	98.0	21	21	AAW10359	Rat oxytocin secre
39	48	98.0	21	21	AAW10366	Human oxytocin sec
40	48	98.0	21	22	AAW62520	Bovine CRH releas
41	48	98.0	21	22	AAW62528	Rat CRH releas
42	48	98.0	21	22	AAW62535	Human CRH releas
43	48	98.0	22	18	AAW31396	Human type G prote
44	48	98.0	22	18	AAW31389	Rat type G protein
45	48	98.0	22	18	AAW31376	Bovine G protein-c

## ALIGNMENTS

RESULT 1	
AAW95178	standard; Protein; 10 AA.
XX	
AC	AAW95178;
XX	
DT	10-MAR-1999 (first entry)
XX	
DE	Murine pituitary-derived ligand polypeptide antigenic epitope.
XX	
KW	Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;
KW	GRB10; UHR-1; modulator; pituitary; central nervous system; pancreas;
KW	tissue; screen; therapeutic; binding; senile dementia; ligand; murine;
KW	Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;
KW	Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;
KW	secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;
KW	gene therapy; transgenic animal; epitope.
XX	
OS	Mus sp.
XX	
PN	WO9849295-A1.
XX	
PD	05-NOV-1998.
XX	
PF	27-APR-1998; 98WO-JP01923.
XX	
PR	28-APR-1997; 97JP-0109974.
XX	
PA	(TAKE ) TAKEDA CHEM IND LTD.
XX	
PI	Fukusumi S, Hinuma S;
XX	
DR	WPI; 1999-009423/01.
XX	

PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
XX pancreas, and for drug screening  
PS  
XX Disclosure: Page 26; 206pp; English.

CC The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jacob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutagen are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.  
XX  
XX

SO Sequence 10 AA:

Query Match 98.0%; Score 48; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.0054;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RCIRPVGRF 9  
| | | | | | | | | |  
DB 2 RGLRPVGRF 10

RESULT 2

AAW91400 standard; Peptide; 15 AA.

XX AAW91400;

DT 06-APR-1998 (first entry)

DE Synthetic ligand 19P2-I31 peptide II.

KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent; antigen.

XX Synthetic.

OS Synthetic.

PN WO9724436-A2.

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0058419.

PR 12-AUG-1996; 96JP-0211805.

PA (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M,

PI Kawamata Y, Kitada C;

XX WPI; 1997-363672/33.

DR ligand peptide for G protein-coupled receptor - acts by modulating  
XX function in the central nervous system, pancreas and pituitary gland  
PT  
XX Example 43; Page 151; 258pp; English.

CC This peptide contains the partial C-terminal sequence of the synthetic  
CC ligand polypeptide 19P2-I31 which is capable of binding to a G  
CC protein-coupled receptor protein. This peptide is used as an antigen to  
CC prepare rabbit anti-bovine 19P2-I31 antibodies which are used in binding  
CC assays. Pharmaceutical compositions containing this ligand may be used  
CC as a pituitary function modulator, a central nervous system modulator  
CC or a pancreatic function modulator. This ligand could have specific  
CC applications as a prophylactic or therapeutic agent for dementia,  
CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
CC hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia,  
CC hyperglycidaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
CC and/or oligogalactia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand  
CC affecting activation of the G protein-coupled receptor protein.  
XX  
XX

SO Sequence 15 AA:

Query Match 98.0%; Score 48; DB 18; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RCIRPVGRF 9  
| | | | | | | | | |  
DB 7 RGLRPVGRF 15

RESULT 3

AAW97230 standard; Peptide; 15 AA.

XX AAW97230;

DT 06-MAY-1999 (first entry)

DE C-terminal ligand polypeptide derived antigen.

KW G protein-coupled receptor; GPCR; hypovarianism; gonocyte cacosgenesis;  
KW menopausal syndrome; eutrophic; hypometabolism; lactation; modulation;  
KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
KW acromegaly; Chlari-Frommel syndrome; Argon-z del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;  
KW contraceptive; placental function; choriochorion; hydatid mole;  
KW abortion; unfertilized fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia; prolactin secretion.

XX Synthetic.

OS Synthetic.

PN WO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PR (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

PI

XX WPI: 1999-105614/09.  
DR  
XX  
XX  
PT use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
PS  
XX  
XX Example 43: Page 123; 241pp; English.  
CC AAW97229-31 represent a ligand polypeptide derived fragments used to  
CC produce antibodies. The specification describes an agent for modulating  
CC prolactin secretion which comprises a ligand polypeptide or a salt, for  
CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
CC prolactin secretion can be used for treating or preventing  
CC hypogonadism, gonocyst cacosgenesis, menopausal syndrome, euthyroid or  
CC hypometabolism. They can be used for promoting lactation in a domestic  
CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
CC secretion can be used for treating or preventing pituitary adenomatosis,  
CC brain tumour, emmenlopathy, autoimmune disease, prolactinoma,  
CC infertility, impotence, amenorrhea, galactorrhea, acromegaly,  
CC Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright  
CC syndrome, lymphoma, Sheehan syndrome or dysospermia. The inhibitory  
CC agents can also be used as contraceptives. The agents for modulating  
CC placental function can be used for treating or preventing  
CC choriocarcinoma, hydralid mole, lruption mole, abortion, unfertility  
CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
CC oxytocia.  
XX  
XX  
SQ Sequence 15 AA;

Query Match 98.0%; Score 48; DB 20; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
| | | | | | | | | |  
Db 7 rgirpvgrf 15

RESULT 4  
AAV49293  
ID AAV49293 standard; peptide; 15 AA.  
XX  
AC AAV49293;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
DE 19P2 ligand peptide fragment.  
XX  
XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
KM pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
OS Homo sapiens.  
XX  
XX Key Location/Qualifiers  
FH Modified-site 15  
FT /note="C-terminal amide"  
XX  
XX W0960112-A1.  
XX  
XX 25-NOV-1999.  
XX  
XX 20-MAY-1999; 99MO-JP02650.  
XX  
XX 21-MAY-1998; 98JP-0140293.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Matsumoto H, Kitada C, Hinuma S;  
XX WPI: 2000-039381/03.  
XX

PT New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality  
XX  
XX  
PS Disclosure; Page 26; 73pp; Japanese.  
XX  
XX

CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAV49290-302 represent peptide fragments of the 19P2 ligand.  
XX  
XX

SQ Sequence 15 AA;

Query Match 98.0%; Score 48; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
| | | | | | | | | |  
Db 7 rgirpvgrf 15

RESULT 5  
AAV49295  
ID AAV49295 standard; peptide; 15 AA.  
XX  
AC AAV49295;  
XX  
DT 22-FEB-2000 (first entry)  
XX

DE 19P2 ligand peptide fragment.

XX Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
KM pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX

OS Homo sapiens.

XX Key Location/Qualifiers  
FH Modified-site 1  
FT /note="N-terminal acetylation"  
FT Modified-site 4  
FT /note="acetylated Tyr"  
FT Modified-site 15  
FT /note="C-terminal amide"

XX W0960112-A1.  
XX  
XX 25-NOV-1999.  
XX

XX 20-MAY-1999; 99MO-JP02650.

XX 21-MAY-1998; 98JP-0140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI: 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in

XX studying diseases related to ligand abnormality

XX Disclosure; Page 26; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19P2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function

CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.  
XX  
XX  
SQ Sequence 15 AA;

Query Match 98.0%; Score 48; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RCIRPVGRF 9  
          |||||  
DB 7 RGLRPVGRF 15

RESULT 6

AAY49296  
ID AAY49296 standard; peptide: 15 AA.

AC AAY49296;

DT 22-FEB-2000 (first entry)

DE 19P2 ligand peptide fragment.

XX Monoclonal antibody: 19P2 ligand; diagnosis; prolactin secretion;  
XX pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX  
XX Homo sapiens.

OS

EH Key Location/Qualifiers

FT Modified-site 15 /note= "C-terminal amide"

XX W09960112-A1.

PN 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP02650.

PF 21-MAY-1998; 98JP-0140293.

PR (TAKE ) TAKEDA CHEM IND LTD.

PA Matsumoto H, Kitada C, Hinuma S;

PI WPI: 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
XX studying diseases related to ligand abnormality -

PT Disclosure: Page 27; 73pp; Japanese.

XX The invention provides a monoclonal antibody which has a specific  
XX refection with the part peptide of the C-terminal of 19P2 ligand or its  
XX derivative. The antibodies can be used in diagnosis or to treat or  
XX prevent diseases associated with abnormality in the pituitary function  
XX regulatory mechanism (e.g. promotion of prolactin secretion), central  
XX nervous regulatory mechanism, and pancreatic function regulatory  
XX mechanism. The antibody-based immunoassay can also be applied in  
XX clarifying the physiological functions of the ligand and its derivative.  
XX Sequences AAY49290-302 represent peptide fragments of the 19P2 ligand.  
XX  
XX  
SQ Sequence 15 AA;

Query Match 98.0%; Score 48; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.0082;  
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RCIRPVGRF 9  
          |||||  
DB 7 RGLRPVGRF 15

RESULT 7

AAW31394  
ID AAW31394 standard; peptide: 20 AA.

AC AAW31394;

DT 06-APR-1998 (first entry)

DE Human type G protein-coupled receptor ligand fragment 4.

XX G protein-coupled receptor; ligand binding; pharmaceutical;  
XX modulator; pituitary; central nervous system; pancreas; prophylactic;  
XX therapeutic agent.  
XX  
XX Homo sapiens.

OS

PN W09724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96WO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

DR WPI: 1997-363672/33.

XX N-PSDB; AAV02431.

XX Claim 2; Page 185; 258pp; English.

XX This sequence represents a peptide fragment from a novel human type  
XX ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
XX sequence represented in AAW31390 and is used in an assay to monitor  
XX ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
XX compositions containing this ligand may be used as a pituitary function  
XX modulator, a central nervous system modulator or a pancreatic function  
XX modulator. This ligand could have specific applications as a  
XX prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
XX syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
XX trauma, growth hormone secretory disease, hyper- and polyphagia,  
XX hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
XX hyperprolactinaemia, diabetes, cancer, pancreaticitis, renal disease,  
XX Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
XX transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
XX acute myocardial infarction, infertility, spinocerebellar degeneration,  
XX bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
XX oligosaccharide. Assays can also be developed to screen compounds which are  
XX capable of altering the binding activity of the ligand affecting  
XX activation of the G protein-coupled receptor protein.  
XX  
XX  
SQ Sequence 20 AA;

Query Match 98.0%; Score 48; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9: Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RCIRPVGRF 9

DB 12 rglrpygrf 20

RESULT 8  
AAMW31387  
ID AAMW31387 standard; Peptide: 20 AA.  
XX  
AC AAMW31387;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Rat type G protein-coupled receptor ligand fragment 4.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
XX therapeutic agent.  
XX  
OS Rat sp.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
XX 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
XX WPI: 1997-363672/33.  
DR N-PSDB; AAV02424.  
DR  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2; Page 180; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the  
CC sequence represented in AAMW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA:

Query Match 98.0%; Score 48; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGIIRPYGRF 9  
|||||

DB 12 rglrpygrf 20

RESULT 9  
AAMW31374  
ID AAMW31374 standard; Peptide: 20 AA.  
XX  
AC AAMW31374;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Bovine G protein-coupled receptor ligand peptide fragment 4.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
XX therapeutic agent.  
XX  
OS Bos taurus.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
XX 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
XX WPI: 1997-363672/33.  
DR N-PSDB; AAV02397.  
DR  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2; Page 161; 258pp; English.  
XX  
CC This sequence represents a peptide fragment of a novel bovine pituitary  
CC derived ligand corresponding to amino acid residues 34 to 53 of the  
CC sequence in AAMW31368 and is used in an assay to monitor ligand binding  
CC to the G protein-coupled receptor protein. Pharmaceutical compositions  
CC containing this ligand may be used as a pituitary function modulator, a  
CC central nervous system modulator or a pancreatic function modulator.  
CC This ligand could have specific applications as a prophylactic or  
CC therapeutic agent for dementia, depression, hyperkinetic syndrome,  
CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
CC growth hormone secretory disease, hyper- and polyphagia,  
CC hypercholesterolaemia, hyperglyceridaemia, hyperlipidaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis, acute  
CC myocardial infarction, spinocerebellar degeneration, bone fracture,  
CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility  
CC and/or oligosaccharia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand thus  
CC affecting activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA:

Query Match 98.0%; Score 48; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGIIRPYGRF 9  
|||||

RESULT 10  
AAM97232  
ID AAM97232 standard; peptide: 20 AA.  
XX  
AC AAM97232;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Bovine pituitary-derived ligand polypeptide fragment.  
XX  
KW Bovine pituitary-derived ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.  
XX  
OS Bos sp.  
XX  
FN MO9858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PE 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
DR WPI: 1999-105614/09.  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 136; 241pp; English.  
XX  
CC The present sequence represents a bovine pituitary-derived ligand  
CC fragment. It is used in the course of the invention. The specification  
CC describes an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenorrhoea, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
SQ Sequence 20 AA:  
Query Match 98.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RGIKPVGRF 9  
| | | | | | | | | |  
DB 12 RGIKPVGRF 20

RESULT 11  
AAM97234  
ID AAM97234 standard; peptide: 20 AA.  
XX  
AC AAM97234;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Rat type ligand polypeptide fragment.  
XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.  
XX  
OS Rattus sp.  
XX  
FN MO9858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PE 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
DR WPI: 1999-105614/09.  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 154; 241pp; English.  
XX  
CC The present sequence represents a rat type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenorrhoea, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
SQ Sequence 20 AA:  
Query Match 98.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RGIKPVGRF 9  
| | | | | | | | | |  
DB 12 RGIKPVGRF 20



RESULT 12  
AAW97236  
ID AAW97236 standard; peptide; 20 AA.  
XX  
AC AAW97236;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Human type ligand polypeptide fragment.  
XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumor; emmenopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chlari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; chorlocarcinoma; hydatid mole;  
KW abortion; unthriftly fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.  
XX  
OS Homo sapiens.  
XX  
PN WO9858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98WO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
XX  
DR WPI; 1999-105614/09.  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
PS Claim 3; Page 166; 241pp; English.  
XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumor, emmenopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chlari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC chorlocarcinoma, hydatid mole, abortion, unthriftly fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
SQ Sequence 20 AA;

Query Match 98.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
12 rgirpvgrf 20

RESULT 13  
AAW95191  
ID AAW95191 standard; peptide; 20 AA.  
XX  
AC AAW95191;  
XX  
DT 10-MAR-1999 (first entry)  
XX  
DE Bovine pituitary-derived ligand polypeptide fragment.  
XX  
KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KW GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KW gene therapy; transgenic animal; bovine.  
XX  
OS Bos sp.  
XX  
PN WO9849295-A1.  
XX  
PD 05-NOV-1998.  
XX  
PF 27-APR-1998; 98WO-JP01923.  
XX  
PR 28-APR-1997; 97JP-0109974.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fukusumi S, Hinuma S;  
XX  
DR WPI; 1999-009423/01.  
XX  
PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening  
XX  
PS Example 19; Page 151; 206pp; English.  
XX  
CC The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of the  
CC polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its muten are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. The present sequence represents a bovine genome-derived ligand  
CC polypeptide fragment which is similar to the murine ligand-polypeptide.  
XX  
SQ Sequence 20 AA;

Query Match 98.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9

Db 12 rglrpgvgrf 20

## RESULT 14

AAW95175 standard; Protein: 20 AA.

AAW95175;

10-MAR-1999 (first entry)

Murine pituitary-derived ligand polypeptide antigenic epitope.

Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor; GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas; tissue; screen; therapeutic; binding; senile dementia; ligand; murine; Alzheimer's disease; Parkinson's disease; Huntington's disease; drug; Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone; secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor; gene therapy; transgenic animal; epitope.

Mus sp.

MO9849295-A1.

05-NOV-1998.

27-APR-1998; 98WO-JP01923.

28-APR-1997; 97JP-0109974.

(TAKE ) TAKEDA CHEM IND LTD.

Fukusumi S, Hinuma S;

WPI: 1999-009423/01.

New polypeptide ligand for orphan G protein coupled receptors - used for treating disorders of central nervous system, pituitary and pancreas, and for drug screening

Disclosure: Page 26; 206pp; English.

The invention relates to a murine pituitary-derived ligand polypeptide which is a ligand for the G-protein coupled orphan receptor designated GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing the ligand polypeptide encoding DNA are used to produce a recombinant ligand polypeptide. The ligand polypeptide, and its fragments, modulate function of the pituitary, central nervous system, pancreas and other tissues and can be used to screen for agents that modulate binding of the polypeptide to the receptor; to quantify the amount of receptor in a sample and to raise antibodies. They may also be used therapeutically, e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs; diabetes; schizophrenia; disorders of growth hormone secretion; cancer; rheumatoid arthritis; epilepsy and many others, also to improve post-operative nutritional status and as vasopressor. Transgenic animals carrying the ligand polypeptide encoding DNA or its mutin are used to study the function of the polypeptide-expressing genes, as models of disease, for drug screening and as source of cell lines. The ligand polypeptide DNA is used as a source of probes and primers; to identify related sequences; in receptor-binding assays; for production of Ab and antisera; in drug development; for gene therapy and to develop transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic epitopes which can be used for the preparation of anti-ligand polypeptide antibody.

Sequence 20 AA;

Query Match 98.0%; Score 48; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;

Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RGIKPVGRF 9

Db 12 rglrpgvgrf 20

## RESULT 15

AAAB10350 standard; peptide: 20 AA.

AAAB10350;

24-NOV-2000 (first entry)

Bovine oxytocin secretion promoting peptide SEQ ID NO: 6.

Bovine; oxytocin secretion promoter; G protein-coupled receptor protein; treatment; disease; pain; atonic bleeding; uterine recovery failure; cow; caesarean section; artificial fertilization; galactostasis; goat; pig; veterinary medicine; milk production.

Bos taurus.

WO200038704-A1.

06-JUL-2000.

22-DEC-1999; 99WO-JP07199.

25-DEC-1998; 98JP-0369585.

(TAKE ) TAKEDA CHEM IND LTD.

Matsumoto H, Kitada C, Hinuma S;

WPI: 2000-452298/39.

Physiologically-active polypeptide recognized as ligand by G protein-coupled receptor protein, for promoting secretion of oxytocin, as drugs for diseases relating to oxytocin secretion and in veterinary medicine -

Claim 5; Page 51; 72pp; Japanese.

This invention describes a novel oxytocin secretion-regulating agent which contains a ligand peptide or its salt for the G protein-coupled receptor protein. It is useful in the form of drugs for ameliorating, preventing and treating diseases relating to oxytocin secretion e.g. weak pains and atonic bleeding, before and after expulsion of placenta, uterine recovery failure, caesarean section, stoppage of artificial fertilization or galactostasis and is also applicable in veterinary medicine for promoting milk production in cow, goat and pig. This sequence represents a bovine peptide which acts as an oxytocin secretion promoter.

Sequence 20 AA;

Query Match 98.0%; Score 48; DB 21; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.011;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Oy 1 RGIKPVGRF 9

Db 12 rglrpgvgrf 20

Search completed: September 13, 2002, 09:18:36  
Job time: 501 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:59 ; Search time 136.62 Seconds  
(without alignments)  
1.788 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_12\_21

Perfect score: 49

Sequence: 1 NCIRPVGREFX 10

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 08  
Maximum Match 1008

Listing first 45 summaries

Database: Issued\_Patents\_AA:\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/6CTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/6ackfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	98.0	15	4	US-08-776-971-93
2	48	98.0	20	3	US-09-105-678A-34
3	48	98.0	20	3	US-09-105-678A-40
4	48	98.0	20	3	US-09-105-678A-46
5	48	98.0	20	4	US-08-776-971-8
6	48	98.0	20	4	US-08-776-971-50
7	48	98.0	20	4	US-08-776-971-64
8	48	98.0	20	4	US-08-776-971-98
9	48	98.0	20	4	US-09-421-208-34
10	48	98.0	20	4	US-09-421-208-40
11	48	98.0	20	4	US-09-421-208-46
12	48	98.0	21	3	US-09-105-678A-28
13	48	98.0	21	3	US-09-105-678A-35
14	48	98.0	21	3	US-09-105-678A-41
15	48	98.0	21	3	US-09-105-678A-47
16	48	98.0	21	4	US-08-776-971-9
17	48	98.0	21	4	US-08-776-971-51
18	48	98.0	21	4	US-08-776-971-65
19	48	98.0	21	4	US-09-421-208-28
20	48	98.0	21	4	US-09-421-208-35
21	48	98.0	21	4	US-09-421-208-41
22	48	98.0	21	4	US-09-421-208-47
23	48	98.0	22	3	US-09-105-678A-36
24	48	98.0	22	3	US-09-105-678A-42
25	48	98.0	22	3	US-09-105-678A-48
26	48	98.0	22	4	US-08-776-971-10
27	48	98.0	22	4	US-08-776-971-52

28	48	98.0	22	4	US-08-776-971-66	Sequence 66, Appl
29	48	98.0	22	4	US-08-776-971-73	Sequence 73, Appl
30	48	98.0	22	4	US-09-421-208-36	Sequence 36, Appl
31	48	98.0	22	4	US-09-421-208-42	Sequence 42, Appl
32	48	98.0	22	4	US-09-421-208-48	Sequence 48, Appl
33	48	98.0	31	3	US-09-105-678A-7	Sequence 7, Appl1
34	48	98.0	31	3	US-09-105-678A-8	Sequence 8, Appl1
35	48	98.0	31	3	US-09-105-678A-9	Sequence 9, Appl1
36	48	98.0	31	3	US-09-105-678A-31	Sequence 31, Appl
37	48	98.0	31	3	US-09-105-678A-37	Sequence 37, Appl
38	48	98.0	31	3	US-09-105-678A-43	Sequence 43, Appl
39	48	98.0	31	4	US-09-172-353-4	Sequence 4, Appl1
40	48	98.0	31	4	US-08-776-971-5	Sequence 5, Appl1
41	48	98.0	31	4	US-08-776-971-47	Sequence 47, Appl
42	48	98.0	31	4	US-08-776-971-61	Sequence 61, Appl
43	48	98.0	31	4	US-08-776-971-97	Sequence 97, Appl
44	48	98.0	31	4	US-09-421-208-7	Sequence 7, Appl1
45	48	98.0	31	4	US-09-421-208-8	Sequence 8, Appl1

## ALIGNMENTS

RESULT 1  
US-08-776-971-93  
; Sequence 93, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hinuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESSES:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93  
US-08-776-971-93

Query Match 98.0%: Score 48: DB 4: Length 15;  
Best Local Similarity 100.0%: Pred. No. 0.0024;  
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 RCIRPVGRF 9  
|||||  
DB 7 RCIRPVGRF 15

RESULT 2  
US-09-105-678A-34  
; Sequence 34, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 98.0%: Score 48: DB 3: Length 20;  
Best Local Similarity 100.0%: Pred. No. 0.0032;  
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 RCIRPVGRF 9  
|||||  
DB 12 RCIRPVGRF 20

RESULT 3  
US-09-105-678A-40  
; Sequence 40, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentln Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/105,678A  
; FILING DATE: 26-JUN-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 98.0%: Score 48: DB 3: Length 20;  
Best Local Similarity 100.0%: Pred. No. 0.0032;  
Matches 9: Conservative 0: Mismatches 0: Indels 0: Gaps 0:

OY 1 RCIRPVGRF 9  
|||||  
DB 12 RCIRPVGRF 20

RESULT 4  
US-09-105-678A-46  
; Sequence 46, Application US/09105678A  
; Patent No. 6103882  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 98.0%; Score 48; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||  
DB 12 RGIRPVGRF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||  
DB 12 RGIRPVGRF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||||  
DB 12 RGIRPVGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||||  
DB 12 RGIRPVGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 98:  
US-08-776-971-98

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
DB 12 RGIRPVGRF 20

RESULT 9  
US-09-421-208-34  
Sequence 34, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RGIRPVGRF 9

DB 12 RGIRPVGRF 20

RESULT 10  
US-09-421-208-40  
Sequence 40, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIORITY APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
DB 12 RGIRPVGRF 20

RESULT 11  
US-09-421-208-46  
Sequence 46, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 98.0%; Score 48; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.0032;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGIRPVGRF 9  
DB 12 RGIRPVGRF 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product= "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product= "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product= "Gly-OH or Gly-Arg"  
US-09-105-678A-28

Query Match 98.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGIRPVGRF 9  
DB 12 RGIRPVGRF 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:



TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 98.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||||  
DB 12 RGIRPVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Mishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 98.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||||

DB 12 RGIRPVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Mishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 98.0%; Score 48; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.0034;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||||||  
DB 12 RGIRPVGRF 20

Search completed: September 13, 2002, 09:20:59  
Job time: 624 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:03 ; Search time 172.41 Seconds  
(without alignments)  
5.573 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_12\_21

Perfect score: 49

Sequence: 1 RGIRPYGRFX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_71:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	98.0	83	2	JC7607
2	36	73.5	674	2	G70875
3	35	71.4	272	2	T39533
4	35	71.4	391	1	XXGZAC
5	35	71.4	504	2	B70821
6	35	71.4	519	2	T45447
7	35	71.4	894	2	E82221
8	35	71.4	1056	2	B82557
9	34	69.4	226	2	A65057
10	34	69.4	307	2	S51485
11	34	69.4	328	2	S77236
12	34	69.4	343	2	AH1823
13	33	67.3	290	2	F97646
14	33	67.3	290	2	AD2870
15	33	67.3	313	2	B95351
16	33	67.3	319	2	T36845
17	33	67.3	335	2	E70655
18	33	67.3	353	2	C70502
19	33	67.3	366	2	C70365
20	33	67.3	445	2	T35893
21	33	67.3	477	2	S68367
22	33	67.3	477	2	S65684
23	33	67.3	477	2	A56449
24	33	67.3	485	2	E83663
25	33	67.3	511	2	T40334
26	33	67.3	589	2	G87485
27	33	67.3	1940	2	F75393
28	32	65.3	57	2	I35055
29	32	65.3	149	2	AH2262

30	32	65.3	150	2	S09872	hypothetical prote
31	32	65.3	154	2	T34825	hypothetical prote
32	32	65.3	249	2	S76104	hypothetical prote
33	32	65.3	266	2	T43770	ribosomal protein
34	32	65.3	285	2	T50932	short-chain dehydr
35	32	65.3	292	2	G95926	probable saccharid
36	32	65.3	307	2	C70643	hypothetical prote
37	32	65.3	307	2	AD0883	conserved hypothet
38	32	65.3	374	2	T29118	hypothetical prote
39	32	65.3	309	2	G70947	hypothetical prote
40	32	65.3	388	2	AG3444	transporter BME115
41	32	65.3	409	2	T51126	hypothetical prote
42	32	65.3	410	2	D83573	tRNA nucleotidyl t
43	32	65.3	435	2	A72658	probable isocitrat
44	32	65.3	492	2	G70749	probable hycE prot
45	32	65.3	562	2	S16594	regulatory protein

#### ALIGNMENTS

##### RESULT 1

JC7607 prolactin-releasing peptide - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C:Accession: JC7607

R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishitake, T.; Hosoya, B.; Bloch, Biophys. Res. Commun. 281, 53-56, 2001

A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:

A:Reference number: JC7607; MUID:21092785; PMID:11178959

A:Contents: Spleen

A:Accession: JC7607

A:Molecule type: DNA

A:Residues: 1-83 <YAM>

A:Cross-references: DDBJ:AB040612; DDBJ:AB040613

C:Comments: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PRP

A:Introns: 33/1

Query Match 98.0%; Score 48; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.016;  
Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RGIRPYGRF 9  
DB 44 RGIRPYGRF 52

##### RESULT 2

G70875 probable oxidoreductase (EC 1.5.99.-) - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: G70875

R:COLE, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, J.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holroyd, Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: G70875

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-674 <COL>

A:Cross-references: GB:AL010186; GB:AL123456; NID:g3261493; PIDN:CA15852.1; PID:g269

C:Experimental source: strain H37RV

C:Genetics:

A:Gene: fadH

C:Superfamily: Methylotrophilus methylotrophus W3A1 trimethylamine dehydrogenase  
C:Keywords: 4Fe-4S; metalloprotein; oxidoreductase  
F:337,340,344,356/Binding site: 4Fe-4S cluster (Cys) (covalent) #status predicted

Query Match 73.5%: Score 36; DB 2; Length 674;

Best Local Similarity 85.7%: Pred. No. 36;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GIRPVGR 8  
|:|||||  
Db 492 GVRPVR 498

RESULT 3

hypothetical protein SPBC16A3.19 - fission yeast (Schizosaccharomyces pombe)

C:Species: Schizosaccharomyces pombe

C>Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999

C:Accession: T39533

R:Wood, V.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.

submitted to the EMBL Data Library, February 1998

A:Reference number: Z21862

A:Accession: T39533

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-272 <WOO>

A:Cross-references: EMBL:AL021748; PIDN:CAAI6870.1; GSPDB:GN0067; SPDB:SPBC16A3.19

A:Experimental source: strain 972h-; cosmid c16A3

C:Genetics:

A:Gene: SPDB:SPBC16A3.19

A:Map position: 2

A:Introns: 53/1; 97/3

Query Match 71.4%: Score 35; DB 2; Length 272;

Best Local Similarity 85.7%: Pred. No. 23;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIKRVG 7  
|:|||||  
Db 47 RGLRPVG 53

RESULT 4

XXGAC acetyl-CoA C-acetyltransferase (EC 2.3.1.9) - Zoogloea ramigera

N:Alternate names: acetoacetyl-CoA thiolase; biosynthetic thiolase; thiolase II

C:Species: Zoogloea ramigera

C>Date: 30-Jun-1991 #sequence\_revision 30-Jun-1991 #text\_change 05-May-2000

C:Accession: A26121

R:Peoples, O.P.; Masamune, S.; Walsh, C.T.; Sinskey, A.J.

J. Biol. Chem. 262, 97-102, 1987

A:Title: Biosynthetic thiolase from Zoogloea ramigera. II. Isolation and characterization

A:Reference number: A26121; MUID:87083504

A:Accession: A26121

A:Molecule type: DNA

A:Residues: 1-391 <PEO>

A:Cross-references: EMBL:J02631; NID:9155617; PIDN:AAA27706.1; PID:9155618

A:Experimental source: strain I-16-M, ATCC 19623

C:Comment: The active enzyme, a tetramer of identical chains, catalyzes the reversible r

hesis, and poly-beta-hydroxybutyrate biosynthesis.

C:Genetics:

A:Gene: phbA

C:Superfamily: acetyl-CoA acetyltransferase

C:Keywords: acyltransferase; coenzyme A; homotetramer; ketone body metabolism; poly-beta

F:2-391/Product: acetyl-CoA acetyltransferase #status predicted <MAT>

F:89/Active site: Cys #status predicted

Query Match 71.4%: Score 35; DB 1; Length 391;

Best Local Similarity 75.0%: Pred. No. 33;  
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIKRVG 8  
|:|||||  
Db 266 RGIKRVG 273

RESULT 5

hypothetical protein RV0982 - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: B70821

R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Feldwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulton, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98293987

A:Accession: B70821

A>Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-504 <COL>

A:Cross-references: GB:AL021999; GB:AL123456; NID:93261538; PIDN:CAAI7581.1; PID:9291

A:Experimental source: strain H37RV

C:Genetics:

C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 71.4%: Score 35; DB 2; Length 504;

Best Local Similarity 85.7%: Pred. No. 43;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GIRPVGR 8  
|:|||||  
Db 188 GVRPVR 194

RESULT 6

T45447 probable two-component system sensor [imported] - Mycobacterium leprae

C:Species: Mycobacterium leprae

C>Date: 31-Jan-2000 #sequence\_revision 31-Jan-2000 #text\_change 02-Sep-2000

C:Accession: T45447

R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, February 1998

A:Reference number: Z22967

A:Accession: T45447

A>Status: preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-519 <JAM>

A:Cross-references: EMBL:AL035500; PIDN:CAB36689.1

A:Experimental source: cosmid L373

C:Genetics:

A:Note: MUCB373.27

C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 71.4%: Score 35; DB 2; Length 519;

Best Local Similarity 85.7%: Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GIRPVGR 8  
|:|||||  
Db 188 GVRPVR 194

RESULT 7

E82221 DNA gyrase, chain A VC1258 [imported] - Vibrio cholerae (strain N16961 serogroup O1)

C:Species: Vibrio cholerae

C>Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001

C;Accession: E82221  
R;Heldberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gilm, M.L.; Dodson, R.J.;  
Chardson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, F.  
L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406 477-483, 2000  
A;Title: DNA Sequence of both chromosomes of the cholera pathogen vibrio cholerae.  
A;Reference number: A82035; MUID:20405833  
A;Accession: E82221  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-894 <HEP>  
A;Cross-references: GB:AE004205; GB:AE003852; NID:99655740; PIDN:AAF94417.1; GSPDB:GN001  
A;Experimental source: serogroup O1; strain N16961; biotype El Tor  
C;Genetics:  
A;Gene: VC1258  
A;Map position: 1  
C;Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerase

Query Match 71.4%; Score 35; DB 2; Length 894;  
Best Local Similarity 62.5%; Pred. No. 78;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 GCRPVG 8  
|:|:|:|  
DB 733 GCRPVG 740

RESULT 8  
B82557  
Hypothetical protein XF2445 [Imported] - Xylella fastidiosa (strain 9a5c)  
C;Species: Xylella fastidiosa  
C;Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 20-Aug-2000  
C;Accession: B82557  
R;Anonymous. The Xylella fastidiosa Consortium of the Organization for Nucleotide Sequen  
Nature 406, 151-157, 2000  
A;Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A;Reference number: A82515; MUID:20365717  
A;Note: for a complete list of authors see reference number A59328 below  
A;Accession: B82557  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-1056 <STM>  
A;Cross-references: GB:AE004053; GB:AE003849; NID:99107631; PIDN:AAF85244.1; GSPDB:GN001  
R;Experimental source: strain 9a5c  
R;Simpson, A.J.G.; Rejnach, F.C.; Arruda, P.; Abreu, F.A.; Agencio, M.; Alvarenga, R.; A  
Brlones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carreir, H  
as-Neto, E.; Docena, C.; El-Dorri, H.; Facinani, A.P.; Ferreira, A.J.S.  
submitted to Genbank, June 2000  
A;Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franca, S.C.; Franco, M.C.; Frohm  
J.D.; Junqueira, M.L.; Kemper, E.L.; Klatajima, J.P.; Klieger, J.E.; Kuramae, E.E.; Laiz  
Chado, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins, E  
A;Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C.Y.;  
F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, D.A  
Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.B.; de Sa, R.G.; Santelli, R.V.; Sawasak  
A;Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv  
M.; Tsunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L.; Z  
A;Reference number: A59328  
A;Contents: annotation  
C;Genetics:  
A;Gene: XF2445

Query Match 71.4%; Score 35; DB 2; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 92;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GCRPVG 8  
|:|:|:|  
DB 1047 GCRPVG 1053

RESULT 9

A65057  
Hypothetical protein b2757 - Escherichia coli (strain K-12)  
C;Species: Escherichia coli  
C;Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
C;Accession: A65057  
R;Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
A.; Rose, D.J.; Mau, B.; Shao, Y.  
Science 277, 1453-1462, 1997  
A;Title: The complete genome sequence of Escherichia coli K-12.  
A;Reference number: A64720; MUID:97426617  
A;Accession: A65057  
A;Status: preliminary; nucleic acid sequence not shown; translation not shown  
A;Molecule type: DNA  
A;Residues: 1-226 <BLAT>  
A;Cross-references: GB:AE000359; GB:U00096; NID:91789110; PIDN:AACT5759.1; PID:917891  
A;Experimental source: strain K-12, substrain MG1655  
C;Superfamily: Escherichia coli hypothetical protein b2757

Query Match 69.4%; Score 34; DB 2; Length 226;  
Best Local Similarity 75.0%; Pred. No. 30;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 2 GCRPVG 9  
|:|:|:|  
DB 25 GCRPVG 32

RESULT 10  
S51485  
ribosomal protein S1 - Synechococcus sp. (PCC 6301)  
C;Species: Synechococcus sp.  
A;Variety: PCC 6301  
C;Date: 15-Jul-1995 #sequence\_revision 01-Sep-1995 #text\_change 20-Jun-2000  
C;Accession: S51485  
R;Sugita, M.; Sugita, C.; Sugita, M.  
Mol. Gen. Genet. 246, 142-147, 1995  
A;Title: Structure and expression of the gene encoding ribosomal protein S1 from the  
n CSI.  
A;Reference number: S51483; MUID:95166170  
A;Accession: S51485  
A;Molecule type: DNA  
A;Residues: 1-307 <SDS>  
A;Cross-references: EMBL:D28752; NID:9560122; PIDN:BA05946.1; PID:666973  
A;Experimental source: PCC 6301  
C;Genetics:  
A;Gene: rps1  
C;Superfamily: Synechocystis ribosomal protein S1  
C;Keywords: protein biosynthesis; ribosome

Query Match 69.4%; Score 34; DB 2; Length 307;  
Best Local Similarity 66.7%; Pred. No. 41;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 GCRPVG 9  
|:|:|:|  
DB 205 GCRPVG 213

RESULT 11  
S77236  
ribosomal protein S1 - Synechocystis sp. (strain PCC 6803)  
N;Alternate names: protein sir1356  
C;Species: Synechocystis sp.  
A;Variety: PCC 6803  
C;Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 20-Jun-2000  
C;Accession: S77236  
R;Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O.; K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A;Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.

A:Reference number: S74322; MUID:97061201  
A:Accession: S77236  
A:Status: nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-328 <KUR>  
A:Cross-references: EMBL:DP90907; GB:AB001339; NRD:91652618; PIDN:BAI17570.1; PID:9165265  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996  
C:Genetics:  
A:Gene: rps1  
C:Superfamily: Synecocystis ribosomal protein S1  
C:Keywords: protein biosynthesis; ribosome

Query Match 69.4%; Score 34; DB 2; Length 328;  
Best Local Similarity 66.7%; Pred. No. 44;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||:|:|  
Db 204 RGIRPVGRF 212

RESULT 12  
AH1823  
30S ribosomal protein S1 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH1823  
R:Kanehisa, T.; Nakamura, Y.; Uhlir, C.P.; Kunitz, T.; Sasamoto, S.; Watanabe, A.; Iriyuchi, N.; Shimizu, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. B. 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH1823  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-343 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA07660.1; PID:917135114; GSPDB:GN00179  
C:Genetics:  
A:Experimental source: strain PCC 7120  
A:Gene: rps1  
C:Superfamily: Synecocystis ribosomal protein S1

Query Match 69.4%; Score 34; DB 2; Length 343;  
Best Local Similarity 66.7%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||:|:|  
Db 205 RGIRPVGRF 213

RESULT 13  
P97646  
Probable alpha/beta sulfonates transport permease protein ssuc AGR\_C.4335 [imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 30-Sep-2001 #sequence\_revision 30-Sep-2001 #text\_change 11-Jan-2002  
C:Accession: P97646  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman, A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.; Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tumefaciens  
A:Reference number: A97359; PMID:11743194  
A:Accession: P97646  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE007869; PIDN:AAK8127.1; PID:915157563; GSPDB:GN00169  
C:Genetics:  
A:Gene: AGR\_C.4335  
A:Map position: circular chromosome

Query Match 67.3%; Score 33; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||:|:|  
Db 123 RGLPPIGRNF 131

RESULT 14  
AD2870  
ABC transporter, membrane spanning protein Atu2390 [imported] - Agrobacterium tumefaciens  
C:Species: Agrobacterium tumefaciens  
C:Date: 11-Jan-2002 #sequence\_revision 11-Jan-2002 #text\_change 11-Jan-2002  
C:Accession: AD2870  
R:Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woerger, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavlin, T.; Levy, R.; Li, M.; McClure, K.; Karp, P.; Romero, P.; Zhang, S.  
Science 294, 2317-2323, 2001  
A:Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamster, E.W.  
A:Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.  
A:Reference number: AB2577; PMID:11743193  
A:Accession: AD2870  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-290 <KUR>  
A:Cross-references: GB:AE008688; PIDN:AA43378.1; PID:917740875; GSPDB:GN00186  
A:Experimental source: strain C58 (Dupont)  
C:Genetics:  
A:Gene: Atu2390  
A:Map position: circular chromosome

Query Match 67.3%; Score 33; DB 2; Length 290;  
Best Local Similarity 55.6%; Pred. No. 61;  
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
|||:|:|  
Db 123 RGLPPIGRNF 131

RESULT 15  
B95351  
VirB6 type IV secretion protein [imported] - Sinorhizobium meliloti (strain 1021) mag  
C:Species: Sinorhizobium meliloti  
C:Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: B95351  
R:Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; B...; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.  
Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001  
A:Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti  
A:Reference number: A95262; MUID:21396509; PMID:11481432  
A:Accession: B95351  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313 <KUR>  
A:Cross-references: GB:AE006469; PIDN:AA65372.1; PID:914523833; GSPDB:GN00165  
A:Experimental source: strain 1021, megaplasmid pSymA  
R:Gilbert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau, D.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, L.; Hyman, R.W.; Jones, T.  
A:Title: The composite genome of the legume symbiont Sinorhizobium meliloti.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
C:Contents: annotation  
C:Genetics:  
A:Gene: virB6

A:Genome: plasmid

Query Match 67.3%; Score 33; DB 2; Length 313;  
Best Local Similarity 75.0%; Pred. No. 66;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Oy 1 RGIKPYGR 8  
||:| ||  
Db 297 RGVRAVGR 304

Search completed: September 13, 2002, 09:24:03  
Job time: 778 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:48 ; Search time 80.21 Seconds  
(without alignments)  
4.827 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_12\_21

Perfect score: 49  
Sequence: 1 RGIRPVGRFX 10

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database: SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	48	98.0	83 1 PRRP_RAT	P81278 rattus norv
2	48	98.0	87 1 PRRP_HUMAN	P81277 homo sapien
3	48	98.0	98 1 PRRP_BOVIN	P81264 bos taurus
4	35	71.4	391 1 THIL_ZOORA	P07097 zocogloea ra
5	35	71.4	922 1 GYRA_AERSA	P48369 ascheromias s
6	34	69.4	224 1 YGCI_ECOLI	O46998 escherichia
7	34	69.4	306 1 RSL_SYNP6	P46328 synechococc
8	34	69.4	328 1 RSLA_SYNP3	P73530 synechocyst
9	34	69.4	801 1 TFR2_HUMAN	O9up52 homo sapien
10	33	67.3	477 1 PPOX_HUMAN	P50336 homo sapien
11	33	67.3	477 1 PPOX_MOUSE	P51175 mus musculu
12	33	67.3	485 1 SYE_BACHD	O9kaf6 bacillus ha
13	32	65.3	150 1 ULA7_HCVNA	P16828 human cytom
14	32	65.3	249 1 Y361_SYNP3	O55578 synechocyst
15	32	65.3	330 1 AR72_HUMAN	O43488 homo sapien
16	32	65.3	331 1 AR73_HUMAN	O95154 homo sapien
17	32	65.3	676 1 EXL1_HUMAN	O92935 homo sapien
18	31	63.3	130 1 RL3_PIG	O29293 sus scrofa
19	31	63.3	149 1 ENRN_BPT7	P00641 bacterioph
20	31	63.3	207 1 XDO5_SCHPO	O14198 schizosach
21	31	63.3	230 1 PIRF_HAEIN	P43812 haemophilus
22	31	63.3	233 1 PYRF_PASPU	O9cmm1 pseudomonas
23	31	63.3	250 1 LINC_PSEPA	P50197 mus musculu
24	31	63.3	364 1 MK11_MOUSE	O10376 mycobacteri
25	31	63.3	367 1 GCST_MYCTU	O15759 homo sapien
26	31	63.3	372 1 MK11_HUMAN	O9x191 thermotoga
27	31	63.3	413 1 CSD_THEMA	P71809 mycobacteri
28	31	63.3	430 1 PYRC_MYCTU	P32527 saccharomyc
29	31	63.3	433 1 ZUO1_YEAST	P13468 drosophila
30	31	63.3	463 1 K10_DROME	P18172 drosophila
31	31	63.3	625 1 DHGL_DROME	P18172 drosophila
32	31	63.3	625 1 DHGL_DROME	P18172 drosophila
33	31	63.3	631 1 DNKL_THERO	P96133 thermophilic

34	31	63.3	719 1 NRPI_YEAST	P32770 saccharomyc
35	31	63.3	908 1 ACON_BACSU	P09339 bacillus su
36	31	63.3	1056 1 POL_BIV06	P19560 bovine immu
37	31	63.3	1056 1 POL_BIV27	P19561 bovine immu
38	31	63.3	1416 1 BLK_MOUSE	O88700 mus musculu
39	31	63.3	1561 1 SPAP_STRMU	P23504 streptococc
40	31	63.3	1565 1 PAC_STRMU	P11657 streptococc
41	30	61.2	116 1 RL34_HUMAN	P49207 homo sapien
42	30	61.2	179 1 RK6_GUTTH	O46908 quillardia
43	30	61.2	292 1 CNR8_HUMAN	O9uff9 homo sapien
44	30	61.2	368 1 YG3Y_YEAST	P53295 saccharomyc
45	30	61.2	376 1 MPK5_ARATH	O39025 arabidopsis

## ALIGNMENTS

```

RESULT 1
PRRP_RAT      STANDARD:      PRT;      83 AA.
AC P81278;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Proactin-releasing peptide precursor (PRRP) (Proactin-releasing
DE hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-
DE releasing peptide PRRP20].
GN PRH.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Hinuma S.,
RA Kitada C., Masuo Y., Asano T., Matsunoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A prolactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RP TISSUE SPECIFICITY.
RA PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RT "Tissue distribution of prolactin-releasing peptide (PRRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -!- FUNCTION: Stimulates prolactin (PRL) release and regulates the
CC expression of prolactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -!- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC medulla oblongata and hypothalamus.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL collaboration
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sdb.ch/announce/
CC or send an email to license@sdb.ch).
CC -----
DR EMBL: AB015418; BAA29026.1;
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 21
FT PEPTIDE 22 52
FT PEPTIDE 33 52
FT MOD_RES 52 52
FT SEQUENCE 83 AA: 9215 MW: 10075.26 Da

```

Query Match 96.0%; Score 48; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.004;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
 |||||  
 DB 44 RGIRPVGRF 52

## RESULT 2

PRRP\_HUMAN

ID PRRP\_HUMAN STANDARD; PRT: 87 AA.

AC P8127; 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].  
 GN PRH.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
 NC NCBL\_TaxID=9606;

[1] SEQUENCE FROM N.A.

RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RA "A proactin-releasing peptide in the brain."  
 RT Nature 393:272-276(1998).

RL [2] TISSUE SPECIFICITY.  
 RP PubMed=10498338;

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H., Sumino Y., Fujino M.;  
 RA "Tissue distribution of proactin-releasing peptide (PRP) and its receptor."  
 RT Regul. Pept. 83:1-10(1999).

RL -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
 CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AB015419; BAA29027.1; -  
 DR MIM; 602663; -  
 KW Hormone; Amidation; Signal.

FT SIGNAL 1 22 BY SIMILARITY.  
 FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE 34 53 PROACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

Query Match 98.0%; Score 48; DB 1; Length 87;  
 Best Local Similarity 100.0%; Pred. No. 0.0042;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
 |||||  
 DB 45 RGIRPVGRF 53

RESULT 3  
 PRRP\_BOVIN  
 ID PRRP\_BOVIN STANDARD; PRT: 98 AA.

AC P81264; 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].  
 GN PRH.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 NC NCBL\_TaxID=9913;

GN PRH.  
 RC TISSUE=Brain;  
 RX MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M., Kurokawa T., Nishimura O., Onda H., Fujino M.;  
 RA "A proactin-releasing peptide in the brain."  
 RT Nature 393:272-276(1998).

RL -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----  
 DR EMBL; AB015417; BAA29025.1; -  
 KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT SIGNAL 1 22  
 FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP31.  
 FT PEPTIDE 33 53 PROACTIN-RELEASING PEPTIDE PRRP20.  
 FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
 SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 96.0%; Score 48; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 0.0048;  
 Matches 9; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
 |||||  
 DB 45 RGIRPVGRF 53

RESULT 4  
 THIL\_ZOORA  
 ID THIL\_ZOORA STANDARD; PRT: 391 AA.

AC P07097; 01-APR-1988 (Rel. 07, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 01-OCT-1996 (Rel. 34, Last annotation update)  
 DE Acetyl-CoA acetyltransferase (EC 2.3.1.9) (Acetoacetyl-CoA thiolase).  
 GN PHA.

OS Zoogloea ramigera.  
 OC Bacteria; Proteobacteria; beta subdivision; Rhodocyclus group;  
 OC Zoogloea.

OC NCBL\_TaxID=350;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 19623 / I-16-M;

```

RX MEDLINE-87083504; PubMed-2878929;
RA Peoples O.P., Masamune S., Walsh C.T., Sinskey A.J.;
RT "Biosynthetic thiolase from Zoogloea ramifera. III. Isolation and
RT characterization of the structural gene.";
RL J. Biol. Chem. 262:97-102(1987).
RN [2]
RP REVISION TO 130.
RX MEDLINE-89359356; PubMed-2670935;
RA Peoples O.P., Sinskey A.J.;
RT "Poly-beta-hydroxybutyrate biosynthesis in Alcaligenes eutrophus H16.
RT Characterization of the genes encoding beta-ketothiolase and
RT acetoacetyl-CoA reductase.";
RL J. Biol. Chem. 264:15293-15297(1989).
RN [3]
RP MUTAGENESIS OF CYS-377.
RX MEDLINE-91217075; PubMed-1673680;
RA Palmer M.A.J., Differding E., Gamboni R., Williams S.F., Peoples O.P.,
RA Walsh C.T., Sinskey S.J., Masamune S.;
RT "Biosynthetic thiolase from Zoogloea ramifera. Evidence for a
RT mechanism involving Cys-378 as the active site base.";
RL J. Biol. Chem. 266:8369-8375(1991).
CC -1- CATALYTIC ACTIVITY: 2 acetyl-CoA + CoA + acetoacetyl-CoA.
CC -1- PATHWAY: FIRST STEP IN POLY-BETA-HYDROXYBUTYRATE BIOSYNTHESIS.
CC -1- SUBUNIT: HOMOTETRAMER.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE THIOLASE FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: J02631; AAA27706.1; ALT_SEQ.
DR PIR: A26121; XKGZAC.
DR PIR: A27754; A27754.
DR HSSP: P27796; IAFY.
DR InterPro: IPR002155; Thiolase.
DR Pfam: PF00108; thiolase_1.
DR PROSITE: PS00098; thiolase_C_1.
DR PROSITE: PS00099; thiolase_L_1.
DR PROSITE: PS00737; thiolase_2_1.
DR Transferase: Acyltransferase; PHB biosynthesis.
KW TRANSFERASE; Acyltransferase; PHB biosynthesis.
FT INIT_MET 0
FT ACT_SITE 88 88 SUBSTRATE BINDING (BY SIMILARITY).
FT ACT_SITE 377 377 BASE.
FT MUTAGEN 377 377 C->G: LOSS OF ACTIVITY.
SQ SEQUENCE 391 AA; 40342 MW; 6D2351A1BC04EDD CRC64;

```

```

Query Match 71.4%; Score 35; DB 1; Length 391;
Best Local Similarity 75.0%; Pred. No. 9.3;
Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 RGIRPYGR 8
DB 266 KGIQPLGR 273

RESULT 5
ID GYRA_AERSA STANDARD; PRT; 922 AA.
AC P48369;
DT 01-FEB-1996 (Rel. 33, Created)
DT 01-FEB-1996 (Rel. 33, Last sequence update)
DT 01-NOV-1997 (Rel. 35, Last annotation update)
DE DNA gyrase subunit A (EC 5.99.1.3).
GN GYRA.
OS Aeromonas salmonicida.
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;

```

```

OC Aeromonas.
OX NCBI_TaxID=645;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=2148/89;
RA Opegaard H.;
RL Submitted (Oct-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE OF 33-179 FROM N.A.
RC STRAIN=ATCC 14174;
RX MEDLINE-95142596; PubMed-7840589;
RA Opegaard H., Sorum H.;
RT "gyrA mutations in quinolone-resistant isolates of the fish pathogen
RT Aeromonas salmonicida.";
RL Antimicrob. Agents Chemother. 38:2460-2464(1994).
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining
CC of double-stranded DNA.
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA
CC BREAKAGE AND REJOINING; THE B CHAIN CATALYZES ATP HYDROLYSIS. THE
CC ENZYME FORMS AN A2B2 TETRAMER.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: I47978; AAB41037.1; -.
DR EMBL: I42453; AAB87239.1; -.
DR HSSP: P09097; IAB4.
DR InterPro: IPR002205; DNA_topoisomV.
DR Pfam: PF00521; DNA_topoisomV_1.
DR SMART: SM00434; TOP4C_1.
KW Isomerase; Topoisomerase; DNA-binding.
FT ACT_SITE 122 122 DNA CLEAVAGE (BY SIMILARITY).
SQ SEQUENCE 922 AA; 101333 MW; 8894965DC4217077 CRC64;

```

```

Query Match 71.4%; Score 35; DB 1; Length 922;
Best Local Similarity 62.5%; Pred. No. 22;
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```

```

OY 1 RGIRPYGR 8
DB 758 KGYRPMGR 765

RESULT 6
ID YGCI_ECOLI STANDARD; PRT; 224 AA.
AC 046898;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical protein ygcI precursor.
GN YGCI OR B2757.
OS Escherichia coli.
OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
OC Escherichia.
OX NCBI_TaxID=562;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K12 / MG1655;
RX MEDLINE-97426617; PubMed-9278503;
RA Blattner F.R., Plunkett G., III, Bloch C.A., Perna N.T., Burland V.,
RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,

```

RA Mau B., Shao Y.;  
RT "The complete genome sequence of *Escherichia coli* K-12.";  
Science 277:1453-1474(1997).  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; U29579; AAA69267.1; ALT\_INIT.  
DR EMBL; AE000359; AAC75799.1; ALT\_INIT.  
DR Ecogen; EG13116; YGCI.  
KW Hypothetical protein; signal; Complete proteome.  
FT SIGNAL 1 17 POTENTIAL.  
FT CHAIN 18 224 HYPOTHETICAL PROTEIN YGCI.  
SQ SEQUENCE 224 AA; 25209 MW; 1C42CC09B317D68 CRC64;  
  
Query Match 69.4%; Score 34; DB 1; Length 224;  
Best Local Similarity 75.0%; Pred. No. 8.5;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;  
  
QY 2 GIRPYGRF 9  
DB 23 GTRPTGRF 30  
  
RESULT 7  
RS1\_SYNP6 STANDARD; PRT; 306 AA.  
ID RS1\_SYNP6  
AC P46228;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 30S ribosomal protein S1.  
GN RPSA OR RPS1.  
OS *Synechococcus* sp. (strain PCC 6301) (Anacystis nidulans).  
OC Bacteria; Cyanobacteria; Chroococcales; *Synechococcus*.  
NX NCBI\_TaxID=1139;  
[1]  
RN SEQUENCE FROM N.A., AND SEQUENCE OF 1-13.  
RA MEDLINE=9516170; PubMed=7862084;  
RA Sugita M., Sugita C., Sugitara M.;  
RT "Structure and expression of the gene encoding ribosomal protein S1  
RT from the cyanobacterium *Synechococcus* sp. strain PCC 6301: striking  
RT sequence similarity to the chloroplast ribosomal protein CSI.";  
RL Mol. Gen. Genet. 246:142-147(1995).  
-----  
CC -1- FUNCTION: BINDS MRNA.  
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 3 S1 MOTIF DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; D28752; BAA05946.1; -.  
DR HSSP; P05055; ISRO.  
DR InterPro; IPR003029; S1.  
DR Pfam; PF00575; S1; 3.  
DR SMART; SM00316; S1; 3.  
DR PROSITE; PS50126; S1; 3.  
KW Ribosomal protein; Repeat; RNA-binding.  
FT INTT\_MET 0  
FT DOMAIN 31 100 S1 MOTIF 1.  
FT DOMAIN 118 182 S1 MOTIF 2.  
FT DOMAIN 196 264 S1 MOTIF 3.

SQ SEQUENCE 306 AA; 34492 MW; 82DA42EBE69B97CE CRC64;  
  
Query Match 69.4%; Score 34; DB 1; Length 306;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 RGIRPYGRF 9  
DB 204 RGIRPYGAF 212  
  
RESULT 8  
RS1A\_SYNP3 STANDARD; PRT; 328 AA.  
ID RS1A\_SYNP3  
AC P73530;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE 30S ribosomal protein S1 homolog A.  
GN RPS1A OR SLR1356.  
OS *Synechocystis* sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; *Synechocystis*.  
NX NCBI\_TaxID=1148;  
[1]  
RN SEQUENCE FROM N.A.  
RX MEDLINE=97061201; PubMed=8905231;  
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Nario K.,  
RA Okumura S., Shimpo S., Takeuchi C., Wada T., Watanabe A.,  
RA Yamada M., Yasuda M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT *Synechocystis* sp. strain PCC6803. II. Sequence determination of the  
RT entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
-----  
CC -1- FUNCTION: BINDS MRNA.  
CC -1- SIMILARITY: BELONGS TO THE S1P FAMILY OF RIBOSOMAL PROTEINS.  
CC -1- SIMILARITY: CONTAINS 3 S1 MOTIF DOMAINS.  
-----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
-----  
DR EMBL; D90807; BAA17570.1; -.  
DR HSSP; P05055; ISRO.  
DR InterPro; IPR003029; S1.  
DR Pfam; PF00575; S1; 3.  
DR SMART; SM00316; S1; 3.  
DR PROSITE; PS50126; S1; 3.  
KW Ribosomal protein; Repeat; RNA-binding; Complete proteome.  
FT DOMAIN 31 100 S1 MOTIF 1.  
FT DOMAIN 118 182 S1 MOTIF 2.  
FT DOMAIN 196 264 S1 MOTIF 3.  
SQ SEQUENCE 328 AA; 36570 MW; DC3FF9B1E5A40619 CRC64;  
  
Query Match 69.4%; Score 34; DB 1; Length 328;  
Best Local Similarity 66.7%; Pred. No. 12;  
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;  
  
QY 1 RGIRPYGRF 9  
DB 204 RGIRPYGAF 212  
  
RESULT 9  
TFR2\_HUMAN STANDARD; PRT; 801 AA.  
ID TFR2\_HUMAN

AC G09P53: 0754422; G9HA99: G9NK67  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Transferrin receptor protein 2 (TfR2).  
GN TfR2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
OX NCBI\_TaxId:9606;  
RN [1]  
RN [1]  
RP SEQUENCE FROM N.A. (ISOFORM ALPHA AND BETA).  
RC TISSUE-Erythroleukemia, and Myeloid Leukemia cells;  
RX MEDLINE-9340005; PubMed-10409623;  
RA Kawabata H., Tang R., Hirama T., Yuong P.T., Kawano S., Gombart A.F.,  
RA Koefler H.P.;  
RT "Molecular cloning of transferrin receptor 2: a new member of the  
RT transferrin receptor-like family.";  
RL J. Biol. Chem. 274:20826-20832(1999).  
RN [2]  
RN [2]  
RP SEQUENCE FROM N.A. (ISOFORM GAMMA).  
RX PubMed-9799793;  
RA Gloeckner G., Scherer S., Schattevoy R., Boright A.P., Weber J.,  
RA Tsai L.-C., Rosenthal A.;  
RA "Large scale analysis of two regions in human chromosome 7q22:  
RT annotation of 650 kb of genomic sequence around the EPO and CCR11 loci  
RT reveals 17 genes.";  
RL Genome Res. 8:1060-1073(1998).  
RN [3]  
RN [3]  
RP SEQUENCE OF 1-158 AND 370-801 FROM N.A.  
RC TISSUE-Carcinoma, and Embryo;  
RX Tisogai T., Ota T., Hayashi K., Sugiyama T., Otsuki T., Suzuki Y.,  
RA Mashikawa T., Nagai K., Sugano S., Shiratori A., Sudo H.,  
RA Maekawa M., Hosokiri T., Kaku Y., Kohata H., Kondo H., Sugawara M.,  
RA Takahashi M., Chiba Y., Ishida S., Murakawa K., Ono Y., Takiguchi S.,  
RA Matsunabe S., Kimura K., Murakami K., Ishii S., Kawai Y., Saito K.,  
RA Yamamoto J., Wakamatsu A., Nakamura Y., Nagahara K., Masuno Y.,  
RA Niinomiya K., Iwayanagi T.;  
RT "MEDO human cDNA sequencing project.";  
RL submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.  
RN [4]  
RN [4]  
RP DISEASE.  
RX PubMed-10802645;  
RA Camaschella C., Roetto A., Cali A., De Gobbi M., Garozzo G.,  
RA Carella M., Majorano N., Tocarò A., Gasparini P.;  
RT "The gene TfR2 is mutated in a new type of haemochromatosis mapping to  
RL 7q22.";  
RN Met. Genet. 25:14-15(2000).  
RN [5]  
RN [5]  
RP VARIANT HFE3 LYS-172.  
RX PubMed-11313241;  
RA Roetto A., Tocarò A., Piperno A., Piga A., Longo F., Garozzo G.,  
RA Cali A., De Gobbi M., Gasparini P., Camaschella C.;  
RT "New mutations inactivating transferrin 2 in hemochromatosis type  
RN 3.";  
RL Blood 97:2555-2560(2001).  
CC -1- FUNCTION: Mediates cellular uptake of transferrin-bound iron in a  
CC non-iron dependent manner. May be involved in iron metabolism,  
CC hepatocyte function and erythrocyte differentiation.  
CC -1- SUBUNIT: Homodimer.  
CC -1- SUBCELLULAR LOCATION: Type II membrane protein. The beta isoform  
CC lacks the transmembrane domain and is probably intracellular.  
CC -1- ALTERNATIVE PRODUCTS: 3 isoforms; alpha (shown here), beta and  
CC gamma; are produced by alternative splicing.  
CC -1- TISSUE SPECIFICITY: Predominantly expressed in liver. While the  
CC alpha form is also expressed in spleen, lung, muscle, prostate and  
CC peripheral blood mononuclear cells, the beta form is expressed in  
CC all tissues tested, albeit weakly.  
CC -1- DISEASE: Defects in TfR2 are a cause of hereditary hemochromatosis  
CC type III (HFE3). HFE3 is a disorder of iron homeostasis resulting  
CC in iron overload and has a phenotype indistinguishable from that  
CC of hereditary hemochromatosis (HH). HH is characterized by  
CC abnormal intestinal iron absorption and progressive increase of

CC	total body iron, which results in midlife in clinical
CC	complications including cirrhosis, cardiopathy, diabetes,
CC	endocrine dysfunctions, arthropathy, and susceptibility to liver
CC	cancer. Since the disease complications can be effectively
CC	prevented by regular phlebotomies, early diagnosis is most
CC	important to provide a normal life expectancy to the affected
CC	subjects.
CC	-1- MISCELLANEOUS: The variant lys-172 found in hereditary
CC	hemochromatosis type III affects the putative initiation codon of
CC	the beta isoform thus preventing its translation.
CC	-1- SIMILARITY: BELONGS TO PEPTIDASE FAMILY M2B8.
CC	-----
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC	the European Bioinformatics Institute. There are no restrictions on its
CC	use by non-profit institutions as long as its content is in no way
CC	modified and this statement is not removed. Usage by and for commercial
CC	entities requires a license agreement (see <a href="http://www.isb-slb.ch/announce/">http://www.isb-slb.ch/announce/</a>
CC	or send an email to <a href="mailto:license@isb-slb.ch">license@isb-slb.ch</a> ).
CC	-----
DR	EMBL; AF067864; AAD45561.1; -
DR	EMBL; AF053356; AAC78796.1; -
DR	EMBL; AK022002; BAB13951.1; -
DR	EMBL; AK000421; BAA91153.1; ALT_INIT.
DR	MEROPS; M28.973; -.
DR	MIM; 604720; -.
DR	MIM; 604250; -.
DR	InterPro: IPR003137; PA.
DR	Pfam: PF02225; PA: 1.
KW	Transmembrane; Glycoprotein; Receptor; Signal-anchor;
KW	Alternative splicing; Disease mutation.
FT	DOMAIN 1 83
FT	TRANSMEM 84 104
FT	CYTOPLASMIC (POTENTIAL).
FT	SIGNAL-ANCHOR (TYPE-II MEMBRANE PROTEIN)
FT	(POTENTIAL).
FT	EXTRACELLULAR (POTENTIAL).
FT	ENDOCYTOSIS SIGNAL (POTENTIAL).
FT	INTERCHAIN (POTENTIAL).
FT	DISULFID 111 111
FT	N-TERMINAL (POTENTIAL).
FT	CARBOHYD 240 240
FT	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD 339 339
FT	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD 540 540
FT	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	CARBOHYD 754 754
FT	N-LINKED (GLCNAc. . .) (POTENTIAL).
FT	VASAPPLIC 1 171
FT	MISSING (IN ISOFORM BETA).
FT	VASAPPLIC 343 369
FT	MISSING (IN ISOFORM GAMMA).
FT	VARIANT 172 172
FT	M -> K (IN HEE3).
FT	/FTID-VAR.012738.
FT	R -> RIFLSNOV (IN REF. 2).
FT	CONFLICT 712 712
FT	SEQUENCE 801 AA; 88755 MW; D3D3082BA835413A CRC64;
SO	
Query Match	69.4%; Score 34; DB 1; Length 801;
Best Local Similarity	75.0%; Pred. No. 31;
Matches	5; Conservative 1; Mismatches 1; Indels 0; Gaps 0.
QY	1 RGRIPVGR 8
	:
DB	262 RGVDEVGR 269
RESULT 10	
PPOX_HUMAN	STANDARD; PRT; 477 AA.
ID	PPOX_HUMAN
AC	P50336;
DT	01-OCT-1996 (Rel. 34, Created)
DT	01-OCT-1996 (Rel. 34, Last sequence update)
DT	15-DEC-1998 (Rel. 37, Last annotation update)
DE	Protoporphyrinogen oxidase (EC 1.3.3.4) (PPO).
GN	PPOX.
OS	Homo sapiens (Human).
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC	Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
NCBI	taxid=9606;
RN	[1]

RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE-95229621; PubMed-7713909;  
 RA Nishimura K., Takekani S., Inokuchi H.;  
 RT "Cloning of a human cDNA for protoporphyrinogen oxidase by  
 RT complementation in vivo of a hemc mutant of Escherichia coli.";  
 RL J. Biol. Chem. 270:8076-8080(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Placenta;  
 RX MEDLINE-96367087; PubMed-8771201;  
 RA Dalley T.A., Dalley H.A.;  
 RT "Human protoporphyrinogen oxidase: expression, purification, and  
 RT characterization of the cloned enzyme.";  
 RL Protein Sci. 5:98-105(1996).  
 RN [3]  
 RP VARIANT VP ARG-232, AND VARIANT HIS-304.  
 RX MEDLINE-97005368; PubMed-8852667;  
 RA Deybach J.-C., Puy H., Robreau A.-M., Lamoril J., da Silva V.,  
 RA Grandchamp B., Nordmann Y.;  
 RT "Mutations in the protoporphyrinogen oxidase gene in patients with  
 RT variegate porphyria.";  
 RL Hum. Mol. Genet. 5:407-410(1996).  
 RN [4]  
 RP VARIANT VP TRP-59, AND VARIANT CYS-168.  
 RX MEDLINE-96241380; PubMed-8673113;  
 RA Meisner P.N., Dalley T.A., Hift R.J., Ziman M., Corrigan A.V.,  
 RA Roberts A.G., Meisner D.M., Kirsch R.E., Dalley H.A.;  
 RT "A R59W mutation in human protoporphyrinogen oxidase results in  
 RT decreased enzyme activity and is prevalent in South Africans with  
 RT variegate porphyria.";  
 RL Nat. Genet. 13:95-97(1996).  
 RN [5]  
 RP VARIANT VP CYS-152.  
 RX MEDLINE-98434271; PubMed-9763307;  
 RA Frank J., Poh-Fitzpatrick M.B., King L.E. Jr., Christiano A.M.;  
 RT "The genetic basis of 'Scarsdale Gourmet Diet' variegate porphyria: a  
 RT missense mutation in the protoporphyrinogen oxidase gene.";  
 RL Arch. Dermatol. Res. 290:441-445(1998).  
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN  
 CC IX TO FORM PROTOPORPHYRIN IX.  
 CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) - protoporphyrin-  
 CC IX + H(2)O(2)  
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER.  
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMODIMER.  
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE  
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE (BY SIMILARITY).  
 CC -1- TISSUE SPECIFICITY: EXPRESSED IN HEART, BRAIN, PLACENTA, LUNG,  
 CC LIVER, SKELETAL MUSCLE, KIDNEY, AND PANCREAS.  
 CC -1- DISEASE: DEFECTS IN PPOX ARE THE CAUSE OF PORPHYRIA VARIEGATA  
 CC (VP), A DISEASE CHARACTERIZED BY SKIN HYPERPIGMENTATION AND HAIR  
 CC HYPERTRICHOYSIS, ASSOCIATED WITH ACUTE ATTACKS, LIKE THOSE OF ACUTE  
 CC INTERMITTENT PORPHYRIA.  
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: D38537; BAA07538.1; -  
 CC EMBL: U26446; AAA67690.1; -  
 CC DR EMBL: 600923; -  
 CC DR MIM: 176200; -  
 CC DR MIM: 176200; -  
 CC Porphyria biosynthesis; Heme biosynthesis; Oxidoreductase;  
 CC Flavoprotein; FAD; Mitochondrion; Disease mutation; Polymorphism.  
 CC NP BIND 9 14  
 CC FT 59 59 R -> W (IN VP).  
 CC FT VARIANT 59 59 /FTid=VAR\_003686.

FT VARIANT 152 152 R -> C (IN VP).  
 FT FT /FTid=VAR\_003687.  
 FT VARIANT 168 168 R -> C.  
 FT FT /FTid=VAR\_003688.  
 FT VARIANT 232 232 G -> R (IN VP).  
 FT FT /FTid=VAR\_003689.  
 FT VARIANT 304 304 R -> H.  
 FT FT /FTid=VAR\_003690.  
 SQ SEQUENCE 477 AA; 50765 MW; 2444DEAC2E6C33EE CRC64;  
 Query Match 67.38; Score 33; DB 1; Length 477;  
 Best Local Similarity 85.78; Pred. No. 29;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 Oy 1 RGIRPVG 7  
 Db 59 RGIRPVG 65  
 RESULT 11  
 ID PROX\_MOUSE STANDARD; PRT; 477 AA.  
 AC P51175; P97344;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE Protoporphyrinogen oxidase (EC 1.3.3.4) (PRO).  
 GN PROX.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-96132666; PubMed-8554330;  
 RA Dalley T.A., Dalley H.A., Meisner P., Prasad A.R.;  
 RT "Cloning, sequence, and expression of mouse protoporphyrinogen  
 RT oxidase.";  
 RL Arch. Biochem. Biophys. 324:379-384(1995).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-Erythrocyte.  
 RX MEDLINE-95331315; PubMed-7607249;  
 RA Takekani S., Yoshinaga T., Furukawa T., Kohno H., Tokunaga R.,  
 RA Nishimura K., Inokuchi H.;  
 RT "Induction of terminal enzymes for heme biosynthesis during  
 RT differentiation of mouse erythroleukemia cells.";  
 RL Eur. J. Biochem. 230:760-765(1995).  
 RN [3]  
 RP SUBCELLULAR LOCATION.  
 RC TISSUE-Liver.  
 RX MEDLINE-88153682; PubMed-3346226;  
 RA Ferreira G.C., Andrew T.L., Karr S.W., Dalley H.A.;  
 RT "Organization of the terminal two enzymes of the heme biosynthetic  
 RT pathway. Orientation of protoporphyrinogen oxidase and evidence for a  
 RT membrane complex.";  
 RL J. Biol. Chem. 263:3835-3839(1988).  
 CC -1- FUNCTION: CATALYZES THE 6-ELECTRON OXIDATION OF PROTOPORPHYRINOGEN  
 CC IX TO FORM PROTOPORPHYRIN IX.  
 CC -1- CATALYTIC ACTIVITY: Protoporphyrinogen-IX + O(2) - protoporphyrin-  
 CC IX + H(2)O(2).  
 CC -1- COFACTOR: CONTAINS ONE FAD PER HOMODIMER (BY SIMILARITY).  
 CC -1- PATHWAY: PENULTIMATE STEP IN HEME AND PORPHYRIN BIOSYNTHESIS.  
 CC -1- SUBUNIT: HOMODIMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: BOUND TO THE MITOCHONDRIAL INNER MEMBRANE  
 CC WITH ITS ACTIVE SITE FACING THE CYTOSOLIC SIDE.  
 CC -1- INDUCTION: DURING ERYTHROID DIFFERENTIATION.  
 CC -1- SIMILARITY: BELONGS TO THE PROTOPORPHYRINOGEN OXIDASE FAMILY.  
 CC -----  
 CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its

use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: U25114; AAA96003.1; -  
DR EMBL: D45185; BAA08126.1; -  
DR MGD: MGI:104968; Ppox.  
KM Porphylin biosynthesis; Heme biosynthesis; Oxidoreductase;  
KM Flavoprotein; FAD; Mitochondrion.  
FT NP\_BIND 9 14 FAD (POTENTIAL).  
FT CONFLICT 64 64 A -> T (IN REF. 2).  
FT CONFLICT 66 66 A -> P (IN REF. 2).  
FT CONFLICT 108 108 L -> S (IN REF. 2).  
FT CONFLICT 427 427 W -> C (IN REF. 2).  
SQ SEQUENCE 477 AA; 50870 MW; 8CFBA8120728DE6F CRC64;

Query Match 67.3%; Score 33; DB 1; Length 477;  
Best Local Similarity 85.7%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RGIRPVG 7  
Db 59 KGIRPAG 65

RESULT 12  
SYE\_BACHD STANDARD: PRT: 485 AA.  
AC 09KGF6;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE GlutamyL-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase) (GluRS).  
GN GLTX OR BH0109.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=C-125 / JCM 9153;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N., Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kihara S., Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and genomic sequence comparison with Bacillus subtilis.";  
RL Nucleic Acids Res. 28:4317-4331(2000).  
CC -1 CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) = AMP + diphosphate + L-glutamyl-tRNA(Glu).  
CC -1 SUBUNIT: MONOMER (BY SIMILARITY).  
CC -1 SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1 SIMILARITY: BELONGS TO CLASS-I AMINOACYL-tRNA SYNTHETASE FAMILY.  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: AP001507; BAB03828.1; -  
DR InterPro: IPR009242; tRNA-synL\_c.  
DR InterPro: IPR001412; tRNA-synL\_i.  
DR Pfam: PF00749; tRNA-synL\_c; 1.  
DR PRINTS: PRO0987; TRNASYNTHGU.  
DR PROSITE: PS00178; AA\_tRNA\_LIGASE\_L; 1.  
KM Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding; Complete proteome.

FT SITE 11 21 "HIGH" REGION.  
FT SITE 252 256 "KMSKS" REGION.  
FT BINDING 255 255 ATP (BY SIMILARITY).  
SQ SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match 67.3%; Score 33; DB 1; Length 485;  
Best Local Similarity 66.7%; Pred. No. 30;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
Db 149 KGIRPVGRF 157

RESULT 13  
UUA7\_HCMVA STANDARD: PRT: 150 AA.  
ID UUA7\_HCMVA  
AC P16828;  
DT 01-AUG-1990 (Rel. 15, Created)  
DT 01-AUG-1990 (Rel. 15, Last sequence update)  
DT 01-FEB-1991 (Rel. 17, Last annotation update)  
DE Hypothetical protein ULI07.  
GN ULI07.  
OS Human cytomegalovirus (strain AD169).  
OC Viruses; dsDNA viruses, no RNA stage; Herpesviridae;  
OC Betaherpesvirinae; Cytomegalovirus.  
OX NCBI\_TaxID=10360;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA MEDLINE=90269039; PubMed=2161319;  
RA Chee M.S., Bankier A.T., Beck S., Bohnl R., Brown C.M., Cerny R., Hornell T., Hutchinson C.A. III, Kourilsky T., Martignetti J.A., Reddie E., Satchwell S.C., Tomlinson P., Weston K.M., Barrell B.G.;  
RT "Analysis of the protein-coding content of the sequence of human cytomegalovirus strain AD169.";  
RL Curr. Top. Microbiol. Immunol. 154:125-169(1990).  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL: X17403; CAA35343.1; -  
DR PIR: S09872; S09872.  
KM Hypothetical protein.  
FT CARBOHYD 144 144 N-LINKED (GLCNAC... ) (POTENTIAL).  
SQ SEQUENCE 150 AA; 17373 MW; 042707546C7EB878 CRC64;

Query Match 65.3%; Score 32; DB 1; Length 150;  
Best Local Similarity 66.7%; Pred. No. 15;  
Matches 6; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
Db 57 KGORPTGRF 65

RESULT 14  
Y361\_SYNY3 STANDARD: PRT: 249 AA.  
ID Y361\_SYNY3  
AC Q35578;  
DT 15-DEC-1998 (Rel. 37, Created)  
DT 15-DEC-1998 (Rel. 37, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 28.2 kDa protein slr0361.  
GN slr0361.  
OS Synechocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.

OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=96127529; PubMed=8590279;  
RA Kaneo T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,  
RA Sugiyama M., Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
RT *Synechocystis* sp. strain PCC6803. I. Sequence features in the 1 Mb  
RT region from map positions 64% to 92% of the genome.";  
RL DNA Res. 2:153-166(1995).  
CC -1- SIMILARITY: BELONGS TO FAMILY 1 OF PSEUDODIRIDINE SYNTHASES.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; D63999; BAI0082.1; -;  
DR InterPro: IPR000613; Pseudou\_synth.  
DR InterPro: IPR002942; Psi\_RSU.  
DR InterPro: IPR002942; S4.  
DR Pfam; PF00849; Pseudou\_synth\_2; 1.  
DR Pfam; PF01479; S4; 1.  
DR SMART; SM00363; S4; 1.  
DR PROSITE; PS01149; PSI\_RSU; 1.  
KM Hypothetical protein; Complete proteome.  
SQ SEQUENCE 249 AA; 28228 MW; 315F208030180326 CRC64;  
  
Query Match 65.3%; Score 32; DB 1; Length 249;  
Best Local Similarity 62.5%; Pred. No. 24;  
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 RGIRPVGR 8  
DB 100 KGLHPVGR 107  
  
RESULT 15  
AR72\_HUMAN  
ID AR72\_HUMAN STANDARD; PRT; 330 AA.  
AC 043488; 075749;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Aflatoxin B1 aldehyde reductase 1 (EC 1.-.-.-) (AFB1-AR 1)  
DE (Aldo\_ketoreductase 7).  
GN AKR7A2 OR AFBAR OR AKR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=98244807; PubMed=9576847;  
RA Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;  
RT "Molecular cloning, expression and catalytic activity of a human AKR7  
RT member of the aldo-keto reductase superfamily: evidence that the  
RT major 2-carboxybenzaldehyde reductase from human liver is a homologue  
RT of rat aflatoxin B1-aldehyde reductase.";  
RL Biochem. J. 332:21-34(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99040634; PubMed=9823300;  
RA Premi C., Savelyeva L., Perri P., Schwab M.;  
RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-  
RT 1p36.1 in a region frequently altered in human tumor cells.";  
RL Cancer Res. 58:5014-5018(1998).

RN [3]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING  
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOL. COULD BE  
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC  
CC EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE  
CC REDUCTASE.  
CC SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL; AF026947; AAC52104.1; -;  
DR EMBL; Y16675; CAA76347.1; -;  
DR EMBL; AL035413; CAB72321.1; -;  
DR MIM; 603418; -;  
DR InterPro: IPR001395; Aldo\_ket\_red.  
DR Pfam; PF00248; aldo\_ket\_red; 1.  
KM Oxidoreductase.  
FT ACT\_SITE 112 112 HYDROGEN-BOND DONOR (PROBABLE).  
FT CONFLICT 113 113 A -> T (IN REF. 1).  
SQ SEQUENCE 330 AA; 36618 MW; 3BBF7ED0CA4D54 CRC64;

Query Match 65.3%; Score 32; DB 1; Length 330;  
Best Local Similarity 75.0%; Pred. No. 32;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
OY 2 GIRPVGR 9  
DB 215 GKOPVGR 222

Search completed: September 13, 2002, 09:30:48  
Job time: 1138 sec



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:24 ; Search time 311.85 Seconds  
(without alignments)  
5.547 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_12\_21

Perfect score: 49

Sequence: 1 RGIKPYGRFX 10

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*  
1: sp-archaea:\*  
2: sp-bacteria:\*  
3: sp-fungi:\*  
4: sp-human:\*  
5: sp-invertebrate:\*  
6: sp-mammal:\*  
7: sp-mhc:\*  
8: sp-organelle:\*  
9: sp-phage:\*  
10: sp-plant:\*  
11: sp-rodent:\*  
12: sp-virus:\*  
13: sp-vertebrate:\*  
14: sp-unclassified:\*  
15: sp-virus:\*  
16: sp-bacteriap:\*  
17: sp-archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	46	93.9	117	13	Q9W624
2	43	87.8	420	2	Q93L27
3	36	73.5	426	17	Q979S0
4	36	73.5	674	16	Q50431
5	35	71.4	272	3	O42924
6	35	71.4	371	16	Q981V3
7	35	71.4	504	16	O53895
8	35	71.4	519	16	O925C7
9	35	71.4	894	16	Q9KSJ8
10	35	71.4	1000	5	O9VS40
11	35	71.4	1056	16	Q9PAQ0
12	35	71.4	1172	5	O9VU79
13	34	69.4	329	2	O9RJR7
14	34	69.4	529	3	O9C1M6
15	34	69.4	537	4	O75230
16	34	69.4	780	4	O75422

17	34	69.4	801	4	Q9UP52	Q9UP52 homo sapien
18	33	67.3	139	5	Q9N6S1	Q9N6S1 leishmania
19	33	67.3	153	3	Q9R133	Q9R133 bacterioph
20	33	67.3	205	4	Q96TC9	Q96TC9 homo sapien
21	33	67.3	253	2	P72459	P72459 streptomyce
22	33	67.3	313	16	Q92Y28	Q92Y28 rhizobium m
23	33	67.3	319	2	O88067	O88067 streptomyce
24	33	67.3	335	16	P96221	P96221 mycobacteri
25	33	67.3	353	16	O33194	O33194 mycobacteri
26	33	67.3	366	16	O66950	O66950 aquifex aeo
27	33	67.3	445	2	O50531	O50531 streptomyce
28	33	67.3	477	11	Q99WM3	Q99WM3 mus musculu
29	33	67.3	511	3	O59714	O59714 schizosacch
30	33	67.3	589	16	Q9A718	Q9A718 caulobacter
31	33	67.3	883	10	Q9LEG6	Q9LEG6 lycoperdico
32	33	67.3	901	16	O99YPS	O99YPS streptococc
33	33	67.3	1340	16	Q9RUC8	Q9RUC8 delnoccocus
34	33	67.3	2104	5	Q21281	Q21281 caenorhabd1
35	33	67.3	2104	5	Q964N4	Q964N4 caenorhabd1
36	32	65.3	97	8	O9B5K5	O9B5K5 balanus gta
37	32	65.3	154	2	O54139	O54139 streptomyce
38	32	65.3	183	16	Q92KK8	Q92KK8 rhizobium m
39	32	65.3	183	17	Q979Z6	Q979Z6 thermoplasm
40	32	65.3	266	8	Q23888	Q23888 dictyostell
41	32	65.3	278	16	Q92OR1	Q92OR1 rhizobium m
42	32	65.3	285	2	O9X4W7	O9X4W7 pseudomonas
43	32	65.3	292	16	Q92VM1	Q92VM1 rhizobium m
44	32	65.3	299	16	P95060	P95060 mycobacteri
45	32	65.3	309	2	O86546	O86546 streptomyce

## ALIGNMENTS

RESULT	ID	Query Match	Score	DB	Length
1	Q9W624	Best Local Similarity	77.8%	Pred. No. 0.11:	117
	Q9W624	Matches 7; Conservative		Mismatches 2; Indels 0; Gaps 0;	
	Q9W624	1 RGIKPYGRF 9			
	Q9W624	11:11:111			
	Q9W624	67 RGIKPYGRF 75			
	Q9W624	2			
	Q9W624	PRELIMINARY:			
	Q9W624	PRT: 420 AA.			
	Q9W624	01-DEC-2001 (TREMREL. 19, Created)			
	Q9W624	01-DEC-2001 (TREMREL. 19, Last sequence update)			
	Q9W624	01-DEC-2001 (TREMREL. 19, Last annotation update)			
	Q9W624	01-DEC-2001 (TREMREL. 19, Last annotation update)			

DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
CN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CCM3239;  
RA Kormanec J., Blatakova J., Novakova R., Homeroova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
Streptomyces aureofaciens CCM3239."  
RL Submitted (May-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY033994; AAK61719.1; -  
SO SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;

Query Match 87.8%; Score 43; DB 2; Length 420;  
Best Local Similarity 100.0%; Pred. No. 1.7;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIRPYGRF 9  
Db 48 GIRPYGRF 55  
|||||

RESULT 3  
Q97980 PRELIMINARY; PRT; 426 AA.  
AC Q97980:  
DT 01-OCT-2001 (TrEMBLrel. 18, Created)  
DT 01-OCT-2001 (TrEMBLrel. 18, Last sequence update)  
DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE TVG1120468 PROTEIN.  
GN TVG1120468.  
OS Thermoplasma volcanium.  
OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;  
CC Thermoplasma  
OX NCBI\_TaxID=50339;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=GS51 / DSM 4299 / JCM 9571;  
RX MEDLINE=20570466; PubMed=11121031;  
RA Kawashima T., Amano N., Kolke H., Makino S.-I., Higuchi S.,  
RA Kawashima O., Yamaoka K., Yamazaki M., Kaneshiro K., Kawamoto T.,  
RA Nunoshiba T., Yamamoto Y., Aramaki H., Makino K., Suzuki M.;  
RT "Archaeal adaptation to higher temperatures revealed by genomic  
sequence of Thermoplasma volcanium."  
RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
DR EMBL: AP000985; BAB60232.1; -  
DR InterPro: IPR002881; DUF58.  
DR InterPro: IPR002035; VWFA.  
DR Pfam: PF01882; DUF58; 1.  
DR SMART: SM00337; VMA; 1.  
KM Complete proteome.  
SO SEQUENCE 426 AA; 49092 MW; 63967CF711116FAA CRC64;

Query Match 73.5%; Score 36; DB 17; Length 426;  
Best Local Similarity 75.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GIRPYGRF 9  
Db 304 GIRPYGRF 311  
|||||

RESULT 4  
Q50431 PRELIMINARY; PRT; 674 AA.  
AC Q50431:  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TrEMBLrel. 19, Last annotation update)  
DE 2,4-DIENOTYL-COA REDUCTASE.  
GN FADH OR RV1175C OR MTY005.11C.  
OS Mycobacterium tuberculosis.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1773;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=H37RV;  
RX MEDLINE=98295987; PubMed=9634230;  
RA Cole S.T., Brosch R., Parkhill J., Garner T., Churcher C., Harris D.,  
RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
RA Backett K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R., Devlin K., Felwell T., Gentles S., Hamlin N., Holroyd S.,  
RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
RA Ruster S., Seeger K., Skelton S., Squares S., Squares R.,  
RA Sulten J.E., Taylor K., Whitehead S., Barrell B.G.;  
RT "Deciphering the biology of Mycobacterium tuberculosis from the  
complete genome sequence."  
RL Nature 393:537-544(1998).  
DR EMBL: AL010186; CAA15852.1; -  
DR Tuberculist: RV1175c; -  
DR InterPro: IPR001064; Crystallin.  
DR InterPro: IPR001327; FAD\_Pyr\_redox.  
DR InterPro: IPR00205; NAD\_Binding.  
DR InterPro: IPR001155; Oxidored\_FMN.  
DR Pfam: PF00724; Oxidored\_FMN; 1.  
DR Pfam: PF00070; Pyr\_redox; 1.  
DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
KM Complete proteome.  
SO SEQUENCE 674 AA; 72901 MW; 8F478F74ED3E0A5 CRC64;

Query Match 73.5%; Score 36; DB 16; Length 674;  
Best Local Similarity 85.7%; Pred. No. 75;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIRPYGR 8  
Db 492 GIRPYGR 498  
|||||

RESULT 5  
O42924 PRELIMINARY; PRT; 272 AA.  
ID O42924:  
AC O42924:  
DT 01-JAN-1999 (TrEMBLrel. 09, Created)  
DT 01-JAN-1999 (TrEMBLrel. 09, Last sequence update)  
DT 01-JUN-2001 (TrEMBLrel. 17, Last annotation update)  
DE HYPOTHETICAL 30.9 KDA PROTEIN C16A3.19 IN CHROMOSOME II.  
GN SPBC16A3.19.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Wood V., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.;  
RL Submitted (Feb-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AL021748; CAA16870.1; -  
DR InterPro: IPR001005; MYB\_DNA\_bind.  
DR PROSITE: PS00037; MYB\_1; UNKNOWN\_1.  
KM Hypothetical protein.  
SO SEQUENCE 272 AA; 30901 MW; 8FE0EB292117107A8 CRC64;

Query Match 71.4%; Score 35; DB 3; Length 272;  
Best Local Similarity 85.7%; Pred. No. 45;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

```

OY 1 RGIRPVG 7
    ||:||||
DB 47 RGLRPVG 53

RESULT 6
O981V3 PRELIMINARY; PRT; 371 AA.
AC O981V3:
DT 01-OCT-2001 (TREMBlrel. 18, Created)
DT 01-OCT-2001 (TREMBlrel. 18, Last sequence update)
DE 01-OCT-2001 (TREMBlrel. 18, Last annotation update)
DE MLR9222 PROTEIN.
GN MLR9222.
OS Rhizobium loti (Mesorhizobium loti).
OC Plasmid pMLA.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Phyllobacteriaceae; Mesorhizobium.
OX NCBI_TaxID=381;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MAF303099;
RX MEDLINE=21082930; PubMed=11214968;
RA Kaneko T., Nakamura Y., Sato S., Asamizu E., Kato T., Sasamoto S.,
RA Watanabe A., Idegawa K., Ishikawa A., Kawashima K., Kimura T.,
RA Kishida Y., Kiyokawa C., Kohara M., Matsumoto M., Matsuno A.,
RA Mochizuki Y., Nakayama S., Nakazaki N., Shlimpo S., Sugimoto M.,
RA Takeuchi C., Yamada M., Tabata S.;
RT "Complete genome structure of the nitrogen-fixing symbiotic bacterium
RT Mesorhizobium loti."
RL DNA Res. 7:331-338(2000).
DR EMBL: AP003015; BAB54606.1; -.
KW Plasmid; Complete proteome.
SQ SEQUENCE 371 AA; 42385 MW; 98D9874B5F831D1 CRC64;

Query Match 71.4%; Score 35; DB 16; Length 371;
Best Local Similarity 62.5%; Pred. No. 63;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RGIRPVG 8
    ||:||||
DB 179 RGLRPVG 186

RESULT 7
O53895 PRELIMINARY; PRT; 504 AA.
AC O53895;
DT 01-JUN-1998 (TREMBlrel. 06, Created)
DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE TWO-COMPONENT SENSOR.
GN RV0982 OR MTV044.10.
OS Mycobacterium tuberculosis.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacterium.
OX NCBI_TaxID=1773;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=H37RV;
RX MEDLINE=98895987; PubMed=9634220;
RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,
RA Gordon S.V., Eiglmeyer K., Gas S., Barry C.E. III, Tekaia F.,
RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R., Devlin K., Feltham T., Gentles S., Hamlin N., Holroyd S.,
RA Hornsby T., Jagers K., Krogh A., McLean J., Moule S., Murphy L.,
RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,
RA Ruter S., Seeger K., Skelton S., Squares S., Squares R.,
RA Sulten J.E., Taylor K., Whitehead S., Barrett B.G.;
RT "Deciphering the biology of Mycobacterium tuberculosis from the
RT complete genome sequence."

```

```

RL Nature 393:537-544(1998).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL: AL021999; CAA17581.1; -.
DR HSSP: P02933; 1BXD.
DR Tuberculin; RV0982; -.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR004359; His_kin_slg.
DR Pfam: PF00672; HAMP; 1.
DR Pfam: PF02518; HATPase_C; 1.
DR Pfam: PF00512; signal; 1.
DR SMART: SM00304; HAMP; 1.
DR SMART: SM00387; HATPase_C; 1.
DR SMART: SM00388; HisKA; 1.
KW Complete proteome; kinase; phosphorylation; sensory transduction;
KW transase.
SQ SEQUENCE 504 AA; 54407 MW; 26792A71AC432232 CRC64;

Query Match 71.4%; Score 35; DB 16; Length 504;
Best Local Similarity 85.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GIRPVG 8
    ||:||||
DB 188 RGLRPVG 194

RESULT 8
O925G7 PRELIMINARY; PRT; 519 AA.
AC O925G7;
DT 01-MAY-1999 (TREMBlrel. 10, Created)
DT 01-MAY-1999 (TREMBlrel. 10, Last sequence update)
DE 01-DEC-2001 (TREMBlrel. 19, Last annotation update)
DE PUTATIVE TWO-COMPONENT SYSTEM SENSOR KINASE.
GN ML0175 OR MCB373.27.
OS Mycobacterium leprae.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.
OX NCBI_TaxID=1769;
RN [1]
RP SEQUENCE FROM N.A.
RA Harris D., Taylor K.;
RL Submitted (FEB-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=TN;
RX MEDLINE=21128732; PubMed=11234002;
RA Cole S.T., Eiglmeyer K., Parkhill J., James K.D., Thomson N.R.,
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,
RA Davies R.M., Devlin K., Dutfoy S., Feltham T., Fraser A., Hamlin N.,
RA Holroyd S., Hornsby T., Jagers K., Lacroix C., Maclean J., Moule S.,
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,
RA Ruter S., Seeger K., Simon S., Simmonds M., Skelton J., Squares R.,
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,
RA Barrett B.G.;
RT "Massive gene decay in the leprosy bacillus."
RL Nature 409:1007-1011(2001).
CC -1- SIMILARITY: TO PROKARYOTE SENSORY TRANSDUCTION PROTEINS.
DR EMBL: AL035500; CAB36689.1; -.
DR EMBL: AL583917; CAC29683.1; -.
DR HSSP: P02933; 1BXD.
DR Lepoma; ML0175; -.
DR InterPro: IPR004358; BCTRLSENSOR.
DR InterPro: IPR003660; HAMP.
DR InterPro: IPR003594; HATPase_C.
DR InterPro: IPR003661; His_kinA.
DR InterPro: IPR004359; His_kin_slg.
DR Pfam: PF00672; HAMP; 1.
DR Pfam: PF02518; HATPase_C; 1.

```

DR Pfam: PF00512; signal: 1.  
DR PRINTS: PR00344; BCTRSENSOR.  
DR SMART: SM00304; HAMF: 1.  
DR SMART: SM00387; HATPase\_c: 1.  
DR SMART: SM00388; HleKA: 1.  
KW Complete proteome; Kinase; Phosphorylation; Sensory transduction;  
KW Transferrase  
SQ SEQUENCE 519 AA; 56470 MW; C0F2DAAD57364F10 CRC64;

Query Match 71.4%; Score 35; DB 16; Length 519;  
Best Local Similarity 85.7%; Pred. No. 90;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 GIRPVGR 8  
1:|||||  
Db 188 GLRPVGR 194

RESULT 9  
ID Q9K5J8 PRELIMINARY; PRT; 894 AA.  
AC Q9K5J8;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE DNA GTRASE, SUBUNIT A.  
GN VC1258.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrrio.  
OX NCBI\_TaxID=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RA MEDLINE=20406833; PubMed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Esmailova M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Newman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
cholerae";  
RT Nature 406:477-483(2000).  
RL EMBL: AEO04205; AAF94417.1; -.  
DR HSSP: P09097; IAB4.  
DR TIGR: VC1258; -.  
DR InterPro: IPR002205; DNA\_topoisomIV.  
DR Pfam: PF00521; DNA\_topoisomIV.1.  
DR SMART: SM00434; TOP4c: 1.  
KW Complete proteome.  
SQ SEQUENCE 894 AA; 99025 MW; C7EA42056A08F57F CRC64;

Query Match 71.4%; Score 35; DB 16; Length 894;  
Best Local Similarity 62.5%; Pred. No. 1.6e+02;  
Matches 5; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

OY 1 RGIRPVGR 8  
1:|||||  
Db 733 KGVPRMGR 740

RESULT 10  
ID Q9VS40 PRELIMINARY; PRT; 1000 AA.  
AC Q9VS40;  
DT 01-MAY-2000 (TREMBlrel. 13, Created)  
DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
DE CG8608 PROTEIN.  
GN UNC-13-4A OR CG8608.

OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_TaxID=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; PubMed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers J.H.C., Blazer R.G., Champé M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Milos G.L.G.,  
RA Adair J.F., Adayan A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Baller R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Berooz P.V., Bernier B.P., Bhandari D., Bolshakov S.,  
RA Borokova D., Botchan M.R., Bouck J., Brostein P., Brotter P.,  
RA Burlis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,  
RA Cherry J.M., Cavley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablo B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Glöckner A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Idegyes C.,  
RA Jaisli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Moulton G., Misha N.V., Mobarry C., Morris J., Moshirell A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pauley J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson W., Skupski M.P., Smith T.,  
RA Slater E.C., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wasmann D.A., Weinstock G.M., Weisenbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
RT "The genome sequence of Drosophila melanogaster";  
RL Science 287:2185-2195(2000).  
DR EMBL: AEO03559; AAF50589.1; -.  
DR HSSP: P21707; 1BYN.  
DR FlyBase: FBgn0035756; unc-13-4A.  
DR InterPro: IPR000008; C2.  
DR Pfam: PF00168; C2; 1.  
DR PRINTS: PR00360; C2DOMAIN.  
DR SMART: SM00239; C2; 2.  
DR PROSITE: PSS0004; C2\_DOMAIN\_2; 1.  
SQ SEQUENCE 1000 AA; 114286 MW; FDF14F05A55696A2 CRC64;

Query Match 71.4%; Score 35; DB 5; Length 1000;  
Best Local Similarity 66.7%; Pred. No. 1.8e+02;  
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 RGIRPVGRF 9  
1:|||||  
Db 130 RGVRLGRF 138

RESULT 11  
ID Q9PA00 PRELIMINARY; PRT; 1056 AA.  
AC Q9PA00;  
DT 01-OCT-2000 (TREMBlrel. 15, Created)  
DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)

```

DT 01-DEC-2001 (TREMBLrel. 19, last annotation update)
DE HYPOTHEICAL PROTEIN XF2445.
GN XF2445.
OS Xylella fastidiosa.
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;
OC Xylella.
OX NCBI_TaxID=2371;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=9A5C;
RX MEDLINE=20365717; PubMed=10910347;
RA Simpson A.J.G., Reinach F.C., Arruda P., Abreu F.A., Acencio M.,
RA Alvenga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Brites M.R.S.,
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carter H.,
RA Colauto N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,
RA Facinoni A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,
RA Fraga J.S., Franco S.C., Franco M.C., Frohme M., Furlan L.R.,
RA Gardier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,
RA Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,
RA Machado M.A., Madeira A.M.B.N., Madeira H.M.F., Martino C.L.,
RA Marques M.A., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Montelero-Vitorello C.B.,
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,
RA Peloto B.R., Pereira G.A.G., Pereira H.A.Jr., Pesquero J.B.,
RA Quaglin R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,
RA de Rosa V.E.Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A.Jr.,
RA da Silveira J.F., Silvestri M.L.Z., Siqueira W.J., de Souza A.A.,
RA de Souza A.P., Terezi M.F., Truffi D., Tsai S.M., Tsubako M.H.,
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,
RA Zaio M.A., Zaiz W., Melandis J., Setubal J.C.;
RT "The genome sequence of the plant pathogen Xylella fastidiosa.";
RL Nature 406:151-159(2000).
DR EMBL: AE004053; AAF85244.1; -.
DR InterPro: IPR000421; FAS8.C.
DR InterPro: IPR001092; Htt_dlm.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 1056 AA; 118471 MW; 361CF2510BDFC10 CRC64;

Query Match
Best Local Similarity 71.4%; Score 35; DB 16; Length 1056;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 GIRPVGR 8
DB 1047 GIRPVGR 1053
1:|||||

```

```

RA Adams M.D., Celinker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galie R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Mortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Heit J.R., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borokova D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,
RA Burks K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Flosser C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Modarri C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Nazny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nuskern D.R., Paclet J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Relvert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sideri-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svitskas R., Tector C., Turner R., Venter E., Wang A.X., Wang X.,
RA Wang Z.-Y., Wasserman D.A., Weinstock G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
DR EMBL: AE003537; AAF49811.1; -.
DR HSSP: 002846; IANL.
DR Flybase: FBgn0036368; CG10738.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR000719; Euk_pkinase.
DR InterPro: IPR001054; Guanylate_cyclase.
DR Pfam: PF01094; ANF_receptor; 2.
DR Pfam: PF00211; guanylate_cyc; 1.
DR Pfam: PF00069; pkinase; 1.
DR SMART: SM00044; CTCG; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW ATP-binding; Lyase; Transferase.
SQ SEQUENCE 1172 AA; 131992 MW; C6450931F82A6D55 CRC64;

Query Match
Best Local Similarity 71.4%; Score 35; DB 5; Length 1172;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 GIRPVGRF 9
DB 434 GIRPVGRF 441
1:|||||

```

```

RESULT 12
Q9VU79 PRELIMINARY; PRT: 1172 AA.
AC Q9VU79;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE CG10738 PROTEIN.
GN CG10738.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC Ephydriota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydriidae; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-BERKELEY;
RX MEDLINE=20196006; PubMed=10731132;

```

```

RESULT 13
Q9RJR7 PRELIMINARY; PRT: 329 AA.
AC Q9RJR7;
DT 01-MAY-2000 (TREMBLrel. 13, Created)
DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE PUTATIVE ZINC-BINDING OXIDOREDUCTASE.
GN SCF51.18.

```

```

OS Streptomyces coelicolor.
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=1902;
RN [1]
RC SEQUENCE FROM N.A.
RP STRAIN=A3(2);
RA Seeger K.J., Harris D.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA Cerdano A.M., Parkhill J., Barrell B.G., Rajandream M.A.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=A3(2);
RA MEDLINE=97000351; PubMed=8843436;
RA Redenbach M., Kleiser H.M., Denapalte D., Eichner A., Cullum J.,
RA Kinasli H., Hopwood D.A.;
RT "A set of ordered cosmids and a detailed genetic and physical map for
RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";
RL MOL. Microbiol. 21:77-96(1996).
DR EMBL: AL132707; CAB59716.1; -
DR InterPro: IPR002085; Adh_2inc; 1.
DR Pfam: PF00107; adh_2inc; 1.
SQ SEQUENCE 329 AA; 34155 MW; 5DCAD4FB174FD042 CRC64;

```

```

Query Match          69.4%: Score 34; DB 2; Length 329;
Best Local Similarity 87.5%: Pred. No. 88;
Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 RGIRPVGR 8
    |||||
Db 137 RGIRTVGR 144

```

```

RESULT 14
ID 09C1M8      PRELIMINARY;      PRT;      529 AA.
AC 09C1M8:
DT 01-JUN-2001 (TReMBLrel. 17, Created)
DT 01-JUN-2001 (TReMBLrel. 17, Last sequence update)
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)
DE HEAT-INDUCED CATALASE.
OS Pleurotus sajor-caju (Oyster mushroom).
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;
OC Agaricales; Pleurotaceae; Pleurotus.
OX NCBI_TaxID=50053;
RN [1]
RP SEQUENCE FROM N.A.
RA Jeong M.-J., Park S.-C.;
RT "Cloning of the catalase gene from Pleurotus sajor-caju.";
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF286097; AK15159.1; -.
DR HSSP: P15202; 1A4E.
DR InterPro: IPR002226; Catalase.
DR Pfam: PF00199; catalase; 1.
DR PRINTS: PR00067; CATALASE.
DR PRODOM: PD000510; CATALASE.1.
DR PROSITE: PS00437; CATALASE_1; 1.
DR PROSITE: PS00438; CATALASE_2; UNKNOWN_1.
SQ SEQUENCE 529 AA; 59791 MW; A432BE6445B6AC13 CRC64;

```

```

Query Match          69.4%: Score 34; DB 3; Length 529;
Best Local Similarity 71.4%: Pred. No. 1.5e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 3 IRPVGRF 9
    :|||
Db 298 LRPIGRF 304

```

```

RESULT 15
ID 075230      PRELIMINARY;      PRT;      537 AA.
AC 075230:
DT 01-NOV-1998 (TReMBLrel. 08, Created)
DT 01-NOV-1998 (TReMBLrel. 08, Last sequence update)
DT 01-NOV-1998 (TReMBLrel. 08, Last annotation update)
DE H53 GSI (FRAGMENT).
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimmmerly W., Bondoc M., Cheng J., Connolly K.S., Gunning K.M.,
RA Kadner K., Miguel T., Miller C., Plickuck S., Follard M., Rojerski H.,
RA Sudranian S., Martin C.H.;
RT "Sequencing of human chromosome 5.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Rieke D.O.;
RT "Large Scale Sequence Analysis and Annotation with the Sequence
RT Comparison Analysis (SCAN) System.";
RL Submitted (JUN-1998) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC005178; AAC24489.1; -.
FT NON_TER      1      537
SQ SEQUENCE 537 AA; 58017 MW; 2ED384E9DB64BC5 CRC64;

```

```

Query Match          69.4%: Score 34; DB 4; Length 537;
Best Local Similarity 75.0%: Pred. No. 1.5e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

QY 1 RGIRPVGR 8
    |||||
Db 295 RGIRPFGR 302

```

```

Search completed: September 13, 2002, 09:29:24
Job time: 1069 sec

```

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 13, 2002, 09:18:36 ; Search time 399.68 Seconds  
(without alignments)  
2.501 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_13\_21

Perfect score: 44  
Sequence: 1 GIRPGRFX 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 11073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

A.Geneseq\_032802:\*

- 1: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1980.DAT:\*
- 2: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1981.DAT:\*
- 3: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1982.DAT:\*
- 4: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1983.DAT:\*
- 5: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1984.DAT:\*
- 6: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1985.DAT:\*
- 7: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1986.DAT:\*
- 8: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1987.DAT:\*
- 9: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1988.DAT:\*
- 10: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1989.DAT:\*
- 11: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1990.DAT:\*
- 12: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1991.DAT:\*
- 13: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1992.DAT:\*
- 14: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1993.DAT:\*
- 15: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1994.DAT:\*
- 16: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1995.DAT:\*
- 17: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1996.DAT:\*
- 18: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1997.DAT:\*
- 19: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1998.DAT:\*
- 20: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA1999.DAT:\*
- 21: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2000.DAT:\*
- 22: /SIDSI/gcgdata/hold-geneseq/geneseq-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	97.7	8	22	AA646955
2	43	97.7	10	20	AAW95178
3	43	97.7	15	18	AAW31400
4	43	97.7	15	20	AAW97230
5	43	97.7	15	21	AAV43293
6	43	97.7	15	21	AAV43295
7	43	97.7	15	21	AAV43296
8	43	97.7	20	18	AAW31394
9	43	97.7	20	18	AAW31387
10	43	97.7	20	18	AAW31374
11	43	97.7	20	20	AAW97232

12	43	97.7	20	20	AAW97234	Rat type ligand po
13	43	97.7	20	20	AAW97236	Human type ligand
14	43	97.7	20	20	AAW95191	Bovine pituitary-d
15	43	97.7	20	20	AAW95175	Murine pituitary-d
16	43	97.7	20	21	AAW10350	Bovine oxytocin se
17	43	97.7	20	21	AAW10358	Rat oxytocin secre
18	43	97.7	20	21	AAW10365	Human oxytocin sec
19	43	97.7	20	21	AAW10369	Oxytocin secretion
20	43	97.7	20	21	AAW49294	19P2 ligand peptid
21	43	97.7	20	21	AAV49301	19P2 ligand peptid
22	43	97.7	20	21	AAV49302	19P2 ligand peptid
23	43	97.7	20	22	AAW62519	Bovine CRH releas
24	43	97.7	20	22	AAW62527	Rat CRH releas
25	43	97.7	20	22	AAW62534	Human CRH releas
26	43	97.7	20	22	AAW62538	CRH releasing prot
27	43	97.7	20	22	AAW90992	Proactin releas
28	43	97.7	20	22	AAW90996	Proactin releas
29	43	97.7	20	22	AAW90996	Proactin releas
30	43	97.7	20	22	AAW46954	Peptide PRP20 fra
31	43	97.7	21	18	AAW31397	Synthetic G prote
32	43	97.7	21	18	AAW31395	Human type G prote
33	43	97.7	21	18	AAW31388	Rat type G prote
34	43	97.7	21	18	AAW31375	Bovine G protein-c
35	43	97.7	21	20	AAW97227	Partial ligand pol
36	43	97.7	21	20	AAW87616	Mammalian 19P2 lig
37	43	97.7	21	20	AAW95192	Bovine pituitary-d
38	43	97.7	21	21	AAW10351	Bovine oxytocin se
39	43	97.7	21	21	AAW10359	Rat oxytocin secre
40	43	97.7	21	21	AAW10366	Human oxytocin sec
41	43	97.7	21	22	AAW62520	Bovine CRH releas
42	43	97.7	21	22	AAW62528	Rat CRH releas
43	43	97.7	21	22	AAW62535	Human CRH releas
44	43	97.7	22	18	AAW31396	Human type G prote
45	43	97.7	22	18	AAW31389	Rat type G prote

#### ALIGNMENTS

RESULT 1

ID AAB46955 standard; Protein: 8 AA.

AC AAB46955;

XX

DT 04-MAY-2001 (first entry)

XX

DE Peptide PRP8 fragment.

XX

KW GPR10; UHR-1; PRP receptor; prolactin-releasing peptide; pain;

KM central nervous system disorder; autonomic regulation; analgesic;

KW hypotensive; blood pressure.

XX

OS Unidentified.

XX

FT

FT Modified-site 8 Location/Qualifiers

FT

FT

PN WO200109182-A1.

PD

PD 08-FEB-2001.

XX

PF 03-AUG-2000; 2000MO-F100664.

XX

PR 03-AUG-1999; 99US-0365756.

PR 20-MAR-2000; 2000US-0531567.

XX

PA (JUVA-) JUVANTIA PHARMA LTD OY.

XX

PI Panula PAJ, Pertovaara A, Kalso E, Korpi E;

XX

DR WPI; 2001-182941/18.

XX C-terminal fragments of prolactin-releasing peptide useful for  
PT regulating autonomic functions and in the manufacture of a medicament  
PT for regulating blood pressure -  
XX  
Claim 2, Page 10; 40pp; English  
PS  
PY

PI	Fukusumi S, Hinuma S;
XX	
DR	WPI; 1999-009423/01.
XX	
PT	New polypeptide ligand
PT	for treating disorders

PT New polypeptide ligand for orphan G protein coupled receptors - used for treating disorders of central nervous system, pituitary and pancreas, and for drug screening

PS Disclosure; Page 26; 206pp; English.

CC This invention describes a novel C-terminal fragment (I) of an isolated  
CC prolectin-releasing peptide (PrRP), referred to as PrRP20 and having a  
CC sequence (S1). The invention also describes (1) a therapeutic composition  
CC (C1) comprising (I) or a C-terminal fragment of PrRP referred to as PrRP8  
CC and comprising a sequence (S2); (2) a diagnostic method based on antisera  
CC against PrRP20 for identification of disorders involving the central  
CC nervous system, including those associated with pain or autonomic  
CC regulation, where specific antisera against the N-and/or C-terminal  
CC domains of PrRP is used to identify alterations in PrRP synthesis or  
CC levels; (3) a rat or human receptor encoded by a 1122 nucleotide sequence  
CC (S3), fully defined in the specification; (4) treating a person suffering  
CC from a disorder regulated by a receptor (II) encoded by a sequence of  
CC 1122 nucleotide sequence, fully defined in the specification, located in  
CC the central nervous system, by administering an agonist or antagonist to  
CC the receptor; and (5) treating blood pressure, by blocking of receptors  
CC of PrRP or its C-terminal fragment Glyileargyprovalglyargph-NH<sub>2</sub> (S2).  
CC The products of the invention have analgesic and hypotensive activity.  
CC (I) is useful for regulating autonomic functions, such as increasing a  
CC blood pressure. (I) is useful for treating pain, for manufacturing a  
CC medicament for regulating blood pressure, and for treating pain. Agonist  
CC and antagonist of (II) are useful for treating acute pain, inflammatory  
CC pain and neuropathic pain, for regulating autonomic functions and  
CC treating high blood pressure.

**SQ Sequence 8 AA;**

Query Match	97.78;	Score 43;	DB 22;	Length 8;
-------------	--------	-----------	--------	-----------

Best Local Similarity 100.0%; Pred. No. 6.4e+05,  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy	1	GIRPVGRF	8
Db	1	glrpvgrf	8

RESULT 2

AC AAW95178;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

KM pituitary-derived ligand polypeptide; G-protein-coupled orphan receptor;  
KM GRP94; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor  
KM gene therapy; transgenic animal; epitope.

OS Mus sp.

PN W09849295-A1

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE ) TAKEDA CHEM IND LTD.

XX

PI Fukusumi S, Hinuma S;

DR WPI; 1999-009423/01.

PT New polypeptide ligand for orphan G protein coupled receptors - used for treatment of central nervous system disorders

The invention relates to a murine pituitary-derived ligand polypeptide which is a ligand for the G-protein coupled orphan receptor designated GPR10 (human) or GHR-1 (rat). Cells transformed with a vector containing the ligand polypeptide encoding DNA are used to produce a recombinant ligand polypeptide. The ligand polypeptide, and its fragments, modulate function of the pituitary, central nervous system, pancreas and other tissues and can be used to screen for agents that modulate binding of the polypeptide to the receptor; to quantify the amount of receptor in a sample and to raise antibodies. They may also be used therapeutically, e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's diseases; Cretzfeldt-Jakob disease; poisoning by heavy metals or drugs; diabetes; schizophrenia; disorders of growth hormone secretion; cancer; rheumatoid arthritis, epilepsy and many others. Also to improve post-operative nutritional status and as vasopressor. Transgenic animals carrying the ligand polypeptide encoding DNA or its mutagen are used to study the function of the polypeptide-expressing genes, as models of disease, for drug screening and as source of cell lines. The ligand polypeptide DNA is used as a source of probes and primers; to identify related sequences; in receptor-binding assays; for production of Ab and antisera; in drug development; for gene therapy and to develop transgenic animals. Sequences AA95174 to AA95178 represent antigenic epitopes which can be used for the preparation of anti-ligand polypeptide antibody.

**SQ Sequence 10 AA;**

Query Match	97.78; Score 43; DB 20; Length 10
-------------	-----------------------------------

Best Local Similarity 100.08; Pred. No. 0.029;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy      1 G1RPVGRF 8
        |||||
Db      3 g1RPVGRF 10
```

### RESULT 3

ID AAW31400 standard; Peptide; 15 AA.

AC AAW31400;

DT 06-APR-1998 (first entry)

DE Synthetic ligand 19P2-L31 peptide II.

KM G protein-coupled receptor; ligand binding; pharmaceutical;  
 KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
 KM therapeutic agent; antigen.

Synthetic.

PN W09724436-A2

PD 10-JUL-1997.

PF 26-DEC-1996; 96WO-JP03821.

PR 18-SEP-1996; 96JP-0246573.

PR 15-MAR-1996; 96JP-0059419.

XX



PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kitada C;  
 XX  
 DR WPI: 1997-363672/33.  
 XX  
 PT Ligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Example 43; Page 151; 258pp; English.  
 XX  
 CC This peptide contains the partial C-terminal sequence of the synthetic  
 CC ligand polypeptide 19P2-L31 which is capable of binding to a G  
 CC protein-coupled receptor protein. This peptide is used as an antigen to  
 CC prepare rabbit anti-bovine 19P2-L31 antibodies which are used in binding  
 CC assays. Pharmacological compositions containing this ligand may be used  
 CC as a pituitary function modulator, a central nervous system modulator  
 CC or a pancreatic function modulator. This ligand could have specific  
 CC applications as a prophylactic or therapeutic agent for dementia,  
 CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
 CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
 CC hyper- and polyphagia, hyperlipidaemia, hypercholesterolaemia,  
 CC hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
 CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
 CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
 CC sclerosis, acute myocardial infarction, infertility, splenocerebellar  
 CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
 CC and/or oligogalacta. Assays can also be developed to screen compounds  
 CC which are capable of altering the binding activity of the ligand  
 CC affecting activation of the G protein-coupled receptor protein.  
 CC  
 SQ Sequence 15 AA:  
 XX  
 QY 1 GIRPVGRF 8  
 | | | | |  
 Db 8 girpvgrf 15  
 XX  
 RESULT 4  
 AAM97230  
 ID AAM97230 standard; Peptide: 15 AA.  
 XX  
 AC AAM97230;  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE C-terminal ligand polypeptide derived antigen.  
 XX  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;  
 KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; chorioncarcinoma; hydatid mole;  
 KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia; prolactin secretion.  
 KW  
 OS Synthetic.  
 XX  
 PN WO9858962-A1.  
 XX  
 PD 30-DEC-1998.  
 XX  
 PF 22-JUN-1998; 98WO-JP02765.  
 XX  
 PR 23-JUN-1997; 97JP-0165437.

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 XX  
 DR WPI: 1999-105614/09.  
 XX  
 PT Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Example 43; Page 123; 241pp; English.  
 XX  
 CC AAM97229-31 represent a ligand polypeptide derived fragments used to  
 CC produce antibodies. The specification describes an agent for modulating  
 CC prolactin secretion which comprises a ligand polypeptide or a salt, for  
 CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
 CC prolactin secretion can be used for treating or preventing  
 CC hypovarianism, gonocyst cacogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,  
 CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,  
 CC infertility, impotence, amenorrhea, galactorrhea, acromegaly,  
 CC Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright  
 CC syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory  
 CC agents can also be used as contraceptives. The agents for modulating  
 CC placental function can be used for treating or preventing  
 CC chorioncarcinoma, hydatid mole, interruption mole, abortion, unthrifty  
 CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
 CC oxytocia.  
 CC  
 SQ Sequence 15 AA:  
 XX  
 QY 1 GIRPVGRF 8  
 | | | | |  
 Db 8 girpvgrf 15  
 XX  
 RESULT 5  
 AAY49293  
 ID AAY49293 standard; peptide: 15 AA.  
 XX  
 AC AAY49293;  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE 19P2 ligand peptide fragment.  
 XX  
 KW Monoclonal antibody; 19P2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
 KW  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT Modified-site 15  
 FT /note="C-terminal amide"  
 XX  
 PN WO9960112-A1.  
 XX  
 PD 25-NOV-1999.  
 XX  
 PF 20-MAY-1999; 99WO-JP02650.  
 XX  
 PR 21-MAY-1998; 98JP-0140293.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.

PI Matsumoto H, Kitada C, Hinuma S;  
 XX WPI; 2000-039381/03.  
 XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 XX  
 PS Disclosure; Page 26; 73pp; Japanese.  
 XX  
 CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
 CC  
 SQ Sequence 15 AA;

Query Match 97.7%; Score 43; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
 |||||  
 Db 8 girpvgrf 15

RESULT 6  
 AAY49295 standard; peptide; 15 AA.  
 XX  
 AC AAY49295;

DT 22-FEB-2000 (first entry)  
 XX  
 DE 19p2 ligand peptide fragment.

XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
 KM pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Modified-site 1 /note- "N-terminal acetylation"

FT Modified-site 4 /note- "acetylated Tyr"

FT Modified-site 15 /note- "C-terminal amide"

XX W09960112-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP02650.

XX 21-MAY-1998; 98JP-0140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 XX  
 PS Disclosure; Page 26; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
 CC  
 SQ Sequence 15 AA;

Query Match 97.7%; Score 43; DB 21; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.045;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
 |||||  
 Db 8 girpvgrf 15

RESULT 7  
 AAY49296 standard; peptide; 15 AA.  
 XX  
 AC AAY49296;

DT 22-FEB-2000 (first entry)

XX 19p2 ligand peptide fragment.

XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
 KM pituitary; regulatory mechanism; central nervous system; pancreatic.

XX Homo sapiens.

OS Key Location/Qualifiers

FT Modified-site 15 /note- "C-terminal amide"

XX W09960112-A1.

XX 25-NOV-1999.

XX 20-MAY-1999; 99WO-JP02650.

XX 21-MAY-1998; 98JP-0140293.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Matsumoto H, Kitada C, Hinuma S;

XX WPI; 2000-039381/03.

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
 PT studying diseases related to ligand abnormality -  
 XX  
 PS Disclosure; Page 27; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
 CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
 CC derivative. The antibodies can be used in diagnosis or to treat or  
 CC prevent diseases associated with abnormality in the pituitary function  
 CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
 CC nervous regulatory mechanism, and pancreatic function regulatory  
 CC mechanism. The antibody-based immunoassay can also be applied in  
 CC clarifying the physiological functions of the ligand and its derivative.  
 CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.

SQ Sequence 15 AA;

Query Match 97.7%; Score 43; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.045;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GIRPVGRF 8  
Db 8 girpvgrf 15

RESULT 8  
AAW31394  
ID AAW31394 standard; Peptide: 20 AA.  
XX  
AC AAW31394;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAW02431.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 165; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC truma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreaticitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 97.7%; Score 43; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GIRPVGRF 8  
Db 13 girpvgrf 20

RESULT 9  
AAW31387  
ID AAW31387 standard; Peptide: 20 AA.  
XX  
AC AAW31387;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Rat type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Rat sp.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAW02424.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 180; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the  
CC sequence represented in AAW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC truma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglyceridaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreaticitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligogalactia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 97.7%; Score 43; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.06;

Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIREVGRF 8  
| | | | | | | |  
Db 13 girevgrf 20

RESULT 10  
AAW31374  
ID AAW31374 standard; Peptide: 20 AA.  
XX  
AC AAW31374;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Bovine G protein-coupled receptor ligand peptide fragment 4.  
XX  
KM G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
XX therapeutic agent.  
XX  
OS Bos taurus.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96MO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kikada C;  
DR WPI: 1997-363672/33.  
DR N-PSDB; AAW02397.  
XX  
XX ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX claim 2; Page 161; 258pp; English.  
XX  
XX This sequence represents a peptide fragment of a novel bovine pituitary  
CC derived ligand corresponding to amino acid residues 34 to 53 of the  
CC sequence in AAW31368 and is used in an assay to monitor ligand binding  
CC to the G protein-coupled receptor protein. Pharmaceutical compositions  
CC containing this ligand may be used as a pituitary function modulator. a  
CC central nervous system modulator or a pancreatic function modulator.  
CC This ligand could have specific applications as a prophylactic or  
CC therapeutic agent for dementia, depression, hyperkinetic syndrome,  
CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
CC growth hormone secretory disease, hyper- and polyphagia.  
CC hypercholesterolemia, hyperglycidaemia, hyperlipidaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, amyotrophic lateral sclerosis, acute  
CC myocardial infarction, spinocerebellar degeneration, bone fracture,  
CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility  
CC and/or oligosaccharia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand thus  
XX affecting activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 97.7%; Score 43; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIREVGRF 8  
| | | | | | | |  
Db 13 girevgrf 20

RESULT 11  
AAW97232  
ID AAW97232 standard; peptide: 20 AA.  
XX  
AC AAW97232;  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Bovine pituitary-derived ligand polypeptide fragment.  
XX  
KM Bovine pituitary-derived ligand; modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KM menopausal syndrome; euthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KM contraceptive; placental function; chorioncarcinoma; hydatid mole;  
KM interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
XX abnormal lipidmetabolism; oxytocia.  
XX  
XX Bos sp.  
XX  
PN W09858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
PF 22-JUN-1998; 98MO-JP02765.  
XX  
PR 23-JUN-1997; 97JP-0165437.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
PI WPI: 1999-105614/09.  
XX  
XX use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
XX claim 3; Page 136; 241pp; English.  
XX  
XX The present sequence represents a bovine pituitary-derived ligand  
CC fragment. It is used in the course of the invention. The specification  
CC describes an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC chorioncarcinoma, hydatid mole, interruption mole, abortion, unfertilized fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
SQ Sequence 20 AA;

Query Match 97.7%; Score 43; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPYGRF 8  
| | | | |  
Db 13 girpygrf 20

## RESULT 12

AAW97234  
ID AAW97234 standard; peptide: 20 AA.

AC AAW97234;

DT 06-MAY-1999 (first entry)

DE Rat type ligand polypeptide fragment.

XX Rat type ligand: modulation; prolactin secretion;

KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;

KW menopausal syndrome; euthyroid; hypometabolism; lactation;

KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;

KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;

KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;

KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KW contraceptive; placental function; choriocarcinoma; hydatid mole;

KW interruption; abortion; unfertility fetus; abnormal saccharometabolism;

XX abnormal lipidmetabolism; oxytocia.

OS Rattus sp.

XX W09858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI: 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin

XX secretion or placental function, e.g. for treating menopausal

XX syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3; Page 154; 241pp; English.

XX The present sequence represents a rat type ligand fragment. It

XX is used in the course of the invention. The specification describes

XX an agent for modulating prolactin secretion which comprises a

XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)

XX protein. The agents for promoting prolactin secretion can be used for

XX treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal

XX syndrome, euthyroid or hypometabolism. They can be used for promoting

XX lactation in a domestic mammal and as an aphrodisiac. The agents for

XX inhibiting prolactin secretion can be used for treating or preventing

XX pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,

XX prolactinoma, infertility, impotence, amenorrhea, galactorrhea,

XX acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,

XX Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.

OY 1 GIRPYGRF 8  
| | | | |  
Db 13 girpygrf 20

## RESULT 13

AAW97236  
ID AAW97236 standard; peptide: 20 AA.

AC AAW97236;

DT 06-MAY-1999 (first entry)

DE Human type ligand polypeptide fragment.

XX Rat type ligand: modulation; prolactin secretion;

KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacosgenesis;

KW menopausal syndrome; euthyroid; hypometabolism; lactation;

KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;

KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;

KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;

KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KW contraceptive; placental function; choriocarcinoma; hydatid mole;

KW interruption; abortion; unfertility fetus; abnormal saccharometabolism;

XX abnormal lipidmetabolism; oxytocia.

OS Homo sapiens.

XX W09858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98WO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX WPI: 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin

XX secretion or placental function, e.g. for treating menopausal

XX syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3; Page 166; 241pp; English.

XX The present sequence represents a human type ligand fragment. It

XX is used in the course of the invention. The specification describes

XX an agent for modulating prolactin secretion which comprises a

XX ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)

XX protein. The agents for promoting prolactin secretion can be used for

XX treating or preventing hypovarianism, gonocyst cacosgenesis, menopausal

XX syndrome, euthyroid or hypometabolism. They can be used for promoting

Query Match 97.7%; Score 43; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
| | | | |  
Db 13 girpvgrf 20

## RESULT 14

AAW95191  
ID AAW95191 standard; peptide; 20 AA.

AC AAW95191;

DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; bovine.

XX Bos sp.

OS WO9849295-A1.

PN 05-NOV-1998.

PD 27-APR-1998; 98WO-JP01923.

PF 28-APR-1997; 97JP-0109974.

PR (TAKE ) TAKEDA CHEM IND LTD.

PA Fukusumi S, Hinuma S;

PI WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

PS Example 19; Page 151; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of the  
CC polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera. In drug development; for gene therapy and to develop transgenic  
CC animals. The present sequence represents a bovine genome-derived ligand  
CC polypeptide fragment which is similar to the murine ligand-polypeptide.

XX Sequence 20 AA;

Query Match 97.7%; Score 43; DB 20; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.06;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
| | | | |  
Db 13 girpvgrf 20

## RESULT 15

AAW95175  
ID AAW95175 standard; protein; 20 AA.

AC AAW95175;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; epitope.

XX Mus sp.

OS WO9849295-A1.

PN 05-NOV-1998.

PD 27-APR-1998; 98WO-JP01923.

PF 28-APR-1997; 97JP-0109974.

PR (TAKE ) TAKEDA CHEM IND LTD.

PA Fukusumi S, Hinuma S;

PI WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

PS Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera. In drug development; for gene therapy and to develop transgenic  
CC animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.

XX Sequence 20 AA;

Query Match 97.7%; Score 43; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred.No. 0.06;  
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GIRPYGRF 8  
 |||||||  
 DB 13 girpygrf 20

Search completed: September 13, 2002, 09:18:36  
 Job time: 501 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:59 ; Search time 136.62 seconds  
(without alignments)  
1.609 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_13\_21

Perfect score: 44

Sequence: 1 GIRPYGRFX 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Issued\_Patents\_AA:\*  
1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCITUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	97.7	15	4	US-08-776-971-93
2	43	97.7	20	3	US-09-105-678A-34
3	43	97.7	20	3	US-09-105-678A-40
4	43	97.7	20	3	US-09-105-678A-46
5	43	97.7	20	4	US-08-776-971-8
6	43	97.7	20	4	US-08-776-971-50
7	43	97.7	20	4	US-08-776-971-64
8	43	97.7	20	4	US-08-776-971-98
9	43	97.7	20	4	US-09-421-208-34
10	43	97.7	20	4	US-09-421-208-46
11	43	97.7	20	4	US-09-105-678A-28
12	43	97.7	21	3	US-09-105-678A-35
13	43	97.7	21	3	US-09-105-678A-41
14	43	97.7	21	3	US-09-105-678A-47
15	43	97.7	21	3	US-08-776-971-9
16	43	97.7	21	4	US-08-776-971-51
17	43	97.7	21	4	US-08-776-971-65
18	43	97.7	21	4	US-09-421-208-38
19	43	97.7	21	4	US-09-421-208-35
20	43	97.7	21	4	US-09-421-208-41
21	43	97.7	21	4	US-09-421-208-47
22	43	97.7	22	3	US-09-105-678A-36
23	43	97.7	22	3	US-09-105-678A-42
24	43	97.7	22	3	US-08-776-971-10
25	43	97.7	22	4	US-08-776-971-52
26	43	97.7	22	4	US-08-776-971-52
27	43	97.7	22	4	US-08-776-971-52

28	43	97.7	22	4	US-08-776-971-66	Sequence 66, Appl
29	43	97.7	22	4	US-08-776-971-73	Sequence 73, Appl
30	43	97.7	22	4	US-09-421-208-36	Sequence 36, Appl
31	43	97.7	22	4	US-09-421-208-42	Sequence 42, Appl
32	43	97.7	22	4	US-09-421-208-48	Sequence 48, Appl
33	43	97.7	31	3	US-09-105-678A-7	Sequence 8, Appl1
34	43	97.7	31	3	US-09-105-678A-8	Sequence 9, Appl1
35	43	97.7	31	3	US-09-105-678A-9	Sequence 31, Appl
36	43	97.7	31	3	US-09-105-678A-31	Sequence 37, Appl
37	43	97.7	31	3	US-09-105-678A-37	Sequence 43, Appl
38	43	97.7	31	3	US-09-105-678A-43	Sequence 4, Appl1
39	43	97.7	31	4	US-08-776-971-5	Sequence 47, Appl
40	43	97.7	31	4	US-08-776-971-47	Sequence 51, Appl
41	43	97.7	31	4	US-08-776-971-61	Sequence 97, Appl
42	43	97.7	31	4	US-08-776-971-61	Sequence 7, Appl1
43	43	97.7	31	4	US-09-421-208-7	Sequence 8, Appl1
44	43	97.7	31	4	US-09-421-208-8	Sequence 8, Appl1
45	43	97.7	31	4	US-09-421-208-8	Sequence 8, Appl1

#### ALIGNMENTS

RESULT 1  
US-08-776-971-93  
Sequence 93, Application US/0876971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habalata, Yugo  
Kawamata, Yuj1  
Hosoya, Masaki  
Fuj11, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: (Unknown)  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/038221  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27, 026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-08-776-971-93

Query Match 97.7%: Score 43; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.012;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
|||||||  
Db 8 GIRPVGRF 15

RESULT 2  
US-09-105-678A-34  
Sequence 34, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 97.7%: Score 43; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
|||||||  
Db 13 GIRPVGRF 20

RESULT 3  
US-09-105-678A-40  
Sequence 40, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 97.7%: Score 43; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
|||||||  
Db 13 GIRPVGRF 20

RESULT 4  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentln Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 97.7%; Score 43; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPYGRF 8  
|||||||  
DB 13 GIRPYGRF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPYGRF 8  
|||||||  
DB 13 GIRPYGRF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
DB 13 GIRPVGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
DB 13 GIRPVGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Hadata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
: INFORMATION FOR SEQ ID NO: 98:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 20 amino acids  
: TYPE: amino acid  
: STRANDEDNESS: single  
: TOPOLOGY: linear  
: SEQUENCE DESCRIPTION: SEQ ID NO: 98:  
US-08-776-971-98

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
|||  
Db 13 GIRPVGRF 20

RESULT 9  
US-09-421-208-34  
: Sequence 34, Application US/09421208  
: Patent No. 6258561  
: GENERAL INFORMATION:  
: APPLICANT: Suenaga, Masato  
: APPLICANT: Moriya, Takeo  
: APPLICANT: Tanaka, Yoko  
: APPLICANT: Nishimura, Osamu  
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
: NUMBER OF SEQUENCES: 52  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
: STREET: 130 Water Street  
: CITY: Boston  
: STATE: MA  
: COUNTRY: USA  
: ZIP: 02109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/421,208  
: FILING DATE:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 09/105,678  
: FILING DATE: 26-JUN-1998  
: APPLICATION NUMBER: JP 172118/1997  
: FILING DATE: 27-JUN-1997  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Conlin, David G.  
: REGISTRATION NUMBER: 27,026  
: REFERENCE/DOCKET NUMBER: 48466-342  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 617-523-3400  
: TELEFAX: 617-523-6440  
: INFORMATION FOR SEQ ID NO: 34:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 20 amino acids  
: TYPE: amino acid  
: STRANDEDNESS:  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8

|||||  
Db 13 GIRPVGRF 20

RESULT 10  
US-09-421-208-40  
: Sequence 40, Application US/09421208  
: Patent No. 6258561  
: GENERAL INFORMATION:  
: APPLICANT: Suenaga, Masato  
: APPLICANT: Moriya, Takeo  
: APPLICANT: Tanaka, Yoko  
: APPLICANT: Nishimura, Osamu  
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
: NUMBER OF SEQUENCES: 52  
: CORRESPONDENCE ADDRESS:  
: ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
: STREET: 130 Water Street  
: CITY: Boston  
: STATE: MA  
: COUNTRY: USA  
: ZIP: 02109  
: COMPUTER READABLE FORM:  
: MEDIUM TYPE: Floppy disk  
: COMPUTER: IBM PC compatible  
: OPERATING SYSTEM: PC-DOS/MS-DOS  
: SOFTWARE: Patentin Release #1.0, Version #1.30  
: CURRENT APPLICATION DATA:  
: APPLICATION NUMBER: US/09/421,208  
: FILING DATE:  
: PRIOR APPLICATION DATA:  
: APPLICATION NUMBER: US 09/105,678  
: FILING DATE: 26-JUN-1998  
: APPLICATION NUMBER: JP 172118/1997  
: FILING DATE: 27-JUN-1997  
: ATTORNEY/AGENT INFORMATION:  
: NAME: Conlin, David G.  
: REGISTRATION NUMBER: 27,026  
: REFERENCE/DOCKET NUMBER: 48466-342  
: TELECOMMUNICATION INFORMATION:  
: TELEPHONE: 617-523-3400  
: TELEFAX: 617-523-6440  
: INFORMATION FOR SEQ ID NO: 40:  
: SEQUENCE CHARACTERISTICS:  
: LENGTH: 20 amino acids  
: TYPE: amino acid  
: STRANDEDNESS:  
: TOPOLOGY: linear  
: MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
|||||  
Db 13 GIRPVGRF 20

RESULT 11  
US-09-421-208-46  
: Sequence 46, Application US/09421208  
: Patent No. 6258561  
: GENERAL INFORMATION:  
: APPLICANT: Suenaga, Masato  
: APPLICANT: Moriya, Takeo  
: APPLICANT: Tanaka, Yoko  
: APPLICANT: Nishimura, Osamu  
: TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
: NUMBER OF SEQUENCES: 52  
: CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 97.7%; Score 43; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
Db 13 GIRPVGRF 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product= "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product= "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product= "Gly-OH or Gly-Arg"  
US-09-105-678A-28

Query Match 97.7%; Score 43; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
Db 13 GIRPVGRF 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 97.7%; Score 43; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
DB 13 GIRPVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105.678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 97.7%; Score 43; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
DB 13 GIRPVGRF 20

DB 13 GIRPVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105.678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 97.7%; Score 43; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.017;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
DB 13 GIRPVGRF 20

Search completed: September 13, 2002, 09:20:59  
Job time: 624 sec

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:03 ; Search time 172.41 Seconds

(without alignments)  
5.016 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_13\_21

Perfect score: 44

Sequence: 1 GIRPVGRFX 9

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR\_71:\*  
1: PIR1:\*  
2: PIR2:\*  
3: PIR3:\*  
4: PIR4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	97.7	83	2	JC7607
2	36	81.8	674	2	G70875
3	35	79.5	504	2	B70821
4	35	79.5	519	2	T45447
5	35	79.5	1036	2	B82557
6	34	77.3	226	2	A65057
7	33	75.0	353	2	C70502
8	33	75.0	366	2	C70365
9	33	75.0	445	2	T35893
10	33	75.0	511	2	T40334
11	33	75.0	894	2	B82221
12	32	72.7	149	2	AH2622
13	32	72.7	285	2	T50932
14	32	72.7	292	2	G95926
15	32	72.7	299	2	C70643
16	32	72.7	307	2	AD0883
17	32	72.7	309	2	T29118
18	32	72.7	374	2	G70947
19	32	72.7	388	2	AG3444
20	32	72.7	409	2	T51126
21	32	72.7	435	2	A72658
22	32	72.7	492	2	G70749
23	32	72.7	1059	2	T20802
24	31	70.5	125	2	B87514
25	31	70.5	129	2	C82313
26	31	70.5	142	2	S60747
27	31	70.5	143	2	AD2174
28	31	70.5	207	2	T38961
29	31	70.5	230	2	B64111

30	31	70.5	286	2	A70667	hypothetical prote
31	31	70.5	308	2	AE1382	dipeptidases homol
32	31	70.5	308	2	AG1751	dipeptidases homol
33	31	70.5	319	2	T36845	probable membrane
34	31	70.5	368	2	B87402	conserved hypotnet
35	31	70.5	403	2	C82228	probable trypsin v
36	31	70.5	403	2	E81246	probable L-sorbose
37	31	70.5	413	2	D72260	amino transferase,
38	31	70.5	413	2	D83310	conserved hypotnet
39	31	70.5	413	2	B86094	probable L-sorbose
40	31	70.5	485	2	E83663	glutamy1-trna synt
41	31	70.5	545	2	A87448	conserved hypotnet
42	31	70.5	612	1	A39019	glucose dehydrogen
43	31	70.5	612	2	B39019	glucose dehydrogen
44	31	70.5	612	2	T36210	conserved hypotnet
45	31	70.5	643	1	T07064	seed biotin-contal

#### ALIGNMENTS

##### RESULT 1

JC7607

Proactin-releasing peptide - rat

C:Species: Rattus norvegicus (Norway rat)

C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001

C:Accession: JC7607

R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibasaki, N.; Hashida, T.; Ishizuka, T.; Hosoya, A.

Biophys. Res. Commun. 281, 53-56, 2001

A:Title: Isolation and characterization of the rat proactin-releasing peptide gene:

A:Reference number: JC7607; MUID:21092785; PMID:11178959

A:Contents: Spleen

A:Accession: JC7607

A:Molecule type: DNA

A:Residues: 1-83 <YAM>

A:Cross-references: DDBJ:AB040612; DDBJ:AB040613

A:Comment: This peptide induces arachidonic acid metabolite release from rat anterior

release, and stimulation of ACTH secretion from the pituitary.

C:Genetics:

A:Gene: PRRP

A:Introns: 33/1

Query Match 97.7% Score 43; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.086;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGRF 8  
DB 45 GIRPVGRF 52

##### RESULT 2

G70875

Probable oxidoreductase (EC 1.5.99.-) - Mycobacterium tuberculosis (strain H37RV)

C:Species: Mycobacterium tuberculosis

C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 20-Jun-2000

C:Accession: G70875

R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon

; Connor, R.; Davies, R.; Devlin, K.; Felwell, T.; Gentles, S.; Hamlin, N.; Holroyd,

Rajandram, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellton, S.; Squares, S.

Nature 393, 537-544, 1998

A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.

A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete geno

A:Reference number: A70500; MUID:98295987

A:Accession: G70875

A:Status: preliminary; nucleic acid sequence not shown; translation not shown

A:Molecule type: DNA

A:Residues: 1-674 <COL>

A:Cross-references: GB:AL010186; GB:AL123456; NID:93261493; PIDN:CAA15852.1; PID:9269

A:Experimental source: strain H37RV

C:Genetics:

A:Gene: fadh

C:Superfamily: Methylotrophilus methylotrophus M3a1 trimethylamine dehydrogenase  
C:Keywords: 4fe-4S; metalloprotein; oxidoreductase  
F:337,340,344,356/Binding site: 4fe-4S cluster (Cys) (covalent) #status predicted

Query Match 81.8%; Score 36; DB 2; Length 674;  
Best Local Similarity 85.7%; Pred. No. 21;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
|:|||||  
DB 492 GVRPVGR 498

RESULT 3  
B70821  
hypothetical protein RV0982 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C>Date: 17-Jul-1998 #sequence\_rev1sion 17-Jul-1998 #text\_change 20-Jun-2000  
C:Accession: B70821  
R:Coile, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Feltham, T.; Gentles, S.; Hamlin, N.; Holtroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S. Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; M0ID:98295987  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-504 <COL>  
A:Cross-references: GB:AL021999; GB:AL123456; NID:g3261538; PIDN:CAAL7581.1; PID:g291694  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV0982  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 79.5%; Score 35; DB 2; Length 504;  
Best Local Similarity 85.7%; Pred. No. 25;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
|:|||||  
DB 188 GLRPVGR 194

RESULT 4  
T45447  
probable two-component system sensor [imported] - Mycobacterium leprae  
C:Species: Mycobacterium leprae  
C>Date: 31-Jan-2000 #sequence\_rev1sion 31-Jan-2000 #text\_change 02-Sep-2000  
C:Accession: T45447  
R:James, K.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A. submitted to the EMBL Data Library, February 1998  
A:Reference number: Z22867  
A:Accession: T45447  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-519 <JAM>  
A:Cross-references: EMBL:AL035500; PIDN:CAB36689.1  
A:Experimental source: cosmid L373  
C:Genetics:  
A:Note: MCB373.27  
C:Superfamily: envz protein; sensor histidine kinase homology

Query Match 79.5%; Score 35; DB 2; Length 519;  
Best Local Similarity 85.7%; Pred. No. 26;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
|:|||||

DB 188 GLRPVGR 194

RESULT 5  
B82557  
hypothetical protein XF2445 [imported] - Xylella fastidiosa (strain 9a5c)  
C:Species: Xylella fastidiosa  
C>Date: 18-Aug-2000 #sequence\_rev1sion 20-Aug-2000 #text\_change 20-Aug-2000  
C:Accession: B82557  
R:anonymous, The Xylella fastidiosa Consortium of the Organization for Nucleotide Seq  
Nature 406, 151-157, 2000  
A:Title: The genome sequence of the plant pathogen Xylella fastidiosa.  
A:Reference number: A82515; M0ID:20355717  
A:Note: for a complete list of authors see reference number A59328 below  
A:Accession: B82557  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1056 <SIM>  
A:Cross-references: GB:AE004053; GB:AE003849; NID:g9107631; PIDN:AAF85244.1; GSPDB:GN  
A:Experimental source: strain 9a5c  
R:Simpson, A.J.G.; Reinach, F.C.; Arruda, P.; Abreu, F.A.; Acencio, M.; Alverenga, R.; Birones, M.R.S.; Bueno, M.R.P.; Camargo, A.A.; Camargo, L.E.A.; Carraro, D.M.; Carrer as-Neio, E.; Docena, C.; El-Dorriy, H.; Facincani, A.P.; Ferreira, A.J.S. submitted to Genbank, June 2000  
A:Authors: Ferreira, V.C.A.; Ferro, J.A.; Fraga, J.S.; Franco, S.C.; Franco, M.C.; Fr J.D.; Junqueira, M.L.; Kemper, E.L.; Kitajima, J.P.; Krieger, J.E.; Kuramae, E.E.; La Chido, M.A.; Madeira, A.M.B.N.; Madeira, H.M.F.; Marino, C.L.; Marques, M.V.; Martins A:Authors: Martins, E.M.F.; Matsukuma, A.Y.; Menck, C.F.M.; Miracca, E.C.; Miyaki, C. F.G.; Nunes, L.R.; Oliveira, M.A.; de Oliveira, M.C.; de Oliveira, R.C.; Palmieri, Rodrigues, V.; Rosa, A.J. de M.; de Rosa Jr., V.E.; de Sa, R.G.; Santelli, R.V.; Sawa A:Authors: da Silva, A.C.R.; da Silva, F.R.; da Silva, A.M.; Silva Jr., W.A.; da Silv M.; Tshunako, M.H.; Vallada, H.; Van Sluys, M.A.; Verjovski-Almeida, S.; Vettore, A.L. A:Reference number: A59328  
A:Contents: annotation  
C:Genetics:  
A:Gene: XF2445

Query Match 79.5%; Score 35; DB 2; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 55;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
|:|||||  
DB 1047 GLRPVGR 1053

RESULT 6  
A65057  
hypothetical protein b2757 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_rev1sion 17-Sep-1997 #text\_change 08-Oct-1999  
C:Accession: A65057  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; A.; Rose, D.J.; Mau, B.; Shao, Y. Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; M0ID:97426617  
A:Accession: A65057  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-226 <BLAT>  
A:Cross-references: GB:AE000359; GB:U00096; NID:g1789110; PIDN:AACT5799.1; PID:g17891 A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypothetical protein b2757

Query Match 77.3%; Score 34; DB 2; Length 226;  
Best Local Similarity 75.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GIRPVGR 8  
|:|||||

Db 25 GTRPTGRF 32

RESULT 7  
C70502  
hypothetical protein RV1692 - Mycobacterium tuberculosis (strain H37RV)  
C:Species: Mycobacterium tuberculosis  
C:Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: C70502  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.; Connor, R.; Davies, R.; Devlin, K.; Fieldwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.; Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skellon, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Squares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome  
A:Reference number: A70500; MUID:98295987  
A:Accession: C70502  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-353 <COL>  
A:Cross-references: GB:298268; GB:AL123456; NID:g3261839; PIDN:CAB10949.1; PID:e332855;  
A:Experimental source: strain H37RV  
C:Genetics:  
A:Gene: RV1692

Query Match 75.0%; Score 33; DB 2; Length 353;  
Best Local Similarity 75.0%; Pred. No. 46;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
|:|:|:|:|  
Db 107 GIRPVRRF 114

RESULT 8  
C70365  
twitching motility protein pili - Aquifex aeolicus  
C:Species: Aquifex aeolicus  
C:Date: 08-May-1998 #sequence\_revision 08-May-1998 #text\_change 24-Oct-2000  
C:Accession: C70365  
R:Decker, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; O'V.  
Nature 392, 353-358, 1998  
A:Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.  
A:Reference number: A70300; MUID:98196666  
A:Accession: C70365  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-366 <AOE>  
A:Cross-references: GB:AE000705; NID:g2983310; PIDN:MAC06903.1; PID:g2983313; GB:AE00065  
A:Experimental source: strain VFS  
C:Genetics:  
A:Gene: pili  
C:Superfamily: twitching motility protein pili

Query Match 75.0%; Score 33; DB 2; Length 366;  
Best Local Similarity 75.0%; Pred. No. 48;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
|:|:|:|:|  
Db 87 GVRGVGRF 94

RESULT 9  
T35893  
FAD-dependent oxidoreductase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T35893  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.

submitted to the EMBL Data Library, November 1997  
A:Reference number: Z21592  
A:Accession: T35893  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-445 <OLI>  
A:Cross-references: EMBL:AL009204; PIDN:CAA15814.1; GSPDB:GN00070; SCOEDB:SC9B10.24c  
C:Genetics:  
A:Gene: SCOEDB:SC9B10.24c  
C:Superfamily: L-gulonolactone oxidase

Query Match 75.0%; Score 33; DB 2; Length 445;  
Best Local Similarity 100.0%; Pred. No. 59;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RPVGRF 8  
|:|:|:|:|  
Db 248 RPVGRF 253

RESULT 10  
T40334  
hypothetical protein SPBC3B8.06 - fission yeast (Schizosaccharomyces pombe)  
C:Species: Schizosaccharomyces pombe  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40334  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.; Pohl, T.  
submitted to the EMBL Data Library, March 1998  
A:Reference number: Z21921  
A:Accession: T40334  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-511 <LYN>  
A:Cross-references: EMBL:AL022244; PIDN:CAA18295.1; GSPDB:GN00067; SPDB:SPBC3B8.06  
A:Experimental source: strain 972h-; cosmid c3B8  
C:Genetics:  
A:Gene: SPDB:SPBC3B8.06  
A:Map position: 2

Query Match 75.0%; Score 33; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred. No. 68;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RPVGRF 8  
|:|:|:|:|  
Db 95 RPVGRF 100

RESULT 11  
E82221  
DNA gyrase, chain A VC1258 [Imported] - Vibrio cholerae (strain N16961 serogroup O1)  
C:Species: Vibrio cholerae  
C:Date: 18-Aug-2000 #sequence\_revision 20-Aug-2000 #text\_change 02-Feb-2001  
C:Accession: E82221  
R:Heidelberg, J.F.; Eisen, J.A.; Nelson, W.C.; Clayton, R.A.; Gwinn, M.L.; Dodson, R.; Charson, D.; Ermolaeva, M.D.; Vamathevan, J.; Bass, S.; Qin, H.; Dragol, I.; Sellers, L.; R.R.; Mekalanos, J.J.; Venter, J.C.; Fraser, C.M.  
Nature 406, 477-483, 2000  
A:Title: DNA Sequence of both chromosomes of the cholera pathogen Vibrio cholerae.  
A:Reference number: AB2035; MUID:20406833  
A:Accession: E82221  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-894 <HEI>  
A:Cross-references: GB:AE004205; GB:AE003852; NID:g9655740; PIDN:AAF94417.1; GSPDB:GN  
C:Genetics:  
A:Gene: VC1258  
A:Map position: 1  
C:Superfamily: DNA topoisomerase (ATP-hydrolyzing) chain A; phage T4 DNA topoisomerases

Query Match 75.0%; Score 33; DB 2; Length 894;  
Best Local Similarity 71.4%; Pred. No. 1.2e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
|:|:|:|  
DB 734 GVRPMGR 740

## RESULT 12

AH2262  
hypothetical protein alr3655 [imported] - Anabaena sp. (strain PCC 7120)  
C:Species: Anabaena sp.  
A:Note: Anabaena sp. (strain PCC 7120) is a synonym of Nostoc sp. strain PCC 7120  
C>Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH2262  
R:Kaneko, T.; Nakamura, Y.; Molk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Itiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete genomic sequence of the filamentous nitrogen-fixing Cyanobacterium *Anabaena*  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2262  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-149 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BAW75354.1; PID:g17132788; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr3655

Query Match 72.7%; Score 32; DB 2; Length 149;  
Best Local Similarity 71.4%; Pred. No. 31;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRPVGRF 8  
|:|:|:|  
DB 6 LRPVGRY 12

## RESULT 13

T50932  
Short-chain dehydrogenase/reductase dltI [imported] - *Pseudomonas ableitaphila*  
C:Species: *Pseudomonas ableitaphila*  
C>Date: 21-Jul-2000 #sequence\_revision 21-Jul-2000 #text\_change 28-Jul-2000  
C:Accession: T50932  
R:Martin, V.J.; Mohr, W.W.  
J. Bacteriol. 181, 2675-2682, 1999  
A:Title: A novel aromatic-ring-hydroxylating dioxygenase from the diterpenoid-degrading  
A:Reference number: 225281; MUID:99235742  
A:Accession: T50932  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-285 <MAR>  
A:Cross-references: EMBL:AF119621; PIDN:RAD21071.1  
A:Experimental source: strain BKM-9; ATCC700689  
C:Genetics:  
A:Gene: dltI  
C:Superfamily: ribitol dehydrogenase; short-chain alcohol dehydrogenase homology

Query Match 72.7%; Score 32; DB 2; Length 285;  
Best Local Similarity 71.4%; Pred. No. 60;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
|:|:|:|  
DB 228 GLRPLGR 234

## RESULT 14

G95926  
probable saccharide deacetylase, slightly similar to NodB protein (EC 3.5.1.-) [imported]  
C:Species: *Sinorhizobium meliloti*  
C>Date: 24-Aug-2001 #sequence\_revision 24-Aug-2001 #text\_change 30-Sep-2001  
C:Accession: G95926  
R:Finan, T.M.; Weidner, S.; Wong, K.; Buhrmester, J.; Chain, P.; Vorholter, F.J.; Her  
Proc. Natl. Acad. Sci. U.S.A. 98, 9889-9894, 2001  
A:Title: The complete sequence of the 1,683-kb *psymb* megaplasmid from the N<sub>2</sub>-fixing  
A:Reference number: A95842; MUID:21396508; PMID:11481431  
A:Accession: G95926  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-292 <KUR>  
A:Cross-references: GB:AL591985; PIDN:CAC49079.1; PID:g15140564; GSPDB:GN00167  
A:Experimental source: strain 1021, megaplasmid *psymb*  
R:Gallibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hub  
pelt, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federapfel, N.A.; Fisher, R.  
L.; Hyman, R.W.; Jones, T.  
Science 293, 668-672, 2001  
A:Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelau  
hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh,  
A:Title: The composite genome of the legume symbiont *Sinorhizobium meliloti*.  
A:Reference number: A96039; MUID:21368234; PMID:11474104  
A:Contents: annotation  
C:Genetics:  
A:Gene: SMD21100  
A:Genome: plasmid  
C:Keywords: hydrolase

Query Match 72.7%; Score 32; DB 2; Length 292;  
Best Local Similarity 100.0%; Pred. No. 61;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVG 6  
|:|:|:|  
DB 140 GIRPVG 145

## RESULT 15

C70643  
hypothetical protein RV0712 - *Mycobacterium tuberculosis* (strain H37Rv)  
C:Species: *Mycobacterium tuberculosis*  
C>Date: 17-Jul-1998 #sequence\_revision 17-Jul-1998 #text\_change 22-Oct-1999  
C:Accession: C70643  
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon  
R.; Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd,  
R.; Knodsen, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.  
Nature 393, 537-544, 1998  
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.  
A:Title: Deciphering the biology of *Mycobacterium tuberculosis* from the complete geno  
A:Reference number: A70500; MUID:98295987  
A:Accession: C70643  
A>Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-299 <COL>  
A:Cross-references: GB:284395; GB:AL123456; NID:g3261698; PIDN:CAB06436.1; PID:e29328  
A:Experimental source: strain H37Rv  
C:Genetics:  
A:Gene: RV0712

Query Match 72.7%; Score 32; DB 2; Length 299;  
Best Local Similarity 75.0%; Pred. No. 63;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
|:|:|:|  
DB 196 GTSVGRF 203

Mon Sep 16 09:13:27 2002

us-09-446-543a-73\_copy\_13\_21.rpr

Page 5

Job time: 778 sec

---

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:48 ; Search time 80.21 Seconds  
(without alignments)  
4.345 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_13\_21

Perfect score: 44  
Sequence: 1 GIRPVGRFX 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	43	97.7	83 1	PRRP_RAT
2	43	97.7	87 1	PRRP_HUMAN
3	43	97.7	98 1	PRRP_BOVIN
4	34	77.3	224 1	YGC1_ECOLI
5	33	75.0	922 1	GYRA_AERSA
6	32	72.7	330 1	AR72_HUMAN
7	32	72.7	331 1	AR73_HUMAN
8	31	70.5	207 1	YDOS_SCHPO
9	31	70.5	230 1	PYRF_HAEIN
10	31	70.5	233 1	PYRF_PASMU
11	31	70.5	250 1	LINC_PSEPA
12	31	70.5	413 1	CSD_THEMA
13	31	70.5	485 1	SYE_BACHD
14	31	70.5	625 1	DHGL_DROME
15	31	70.5	626 1	DHGL_DROPS
16	31	70.5	675 1	EXL1_HUMAN
17	31	70.5	719 1	NRPI_YEAST
18	31	70.5	1561 1	SPAP_STRMU
19	31	70.5	1565 1	PAC_STRMU
20	30	68.2	249 1	Y361_SYNY3
21	30	68.2	292 1	CNTB_HUMAN
22	30	68.2	364 1	MK11_MOUSE
23	30	68.2	372 1	MPKS_ARATH
24	30	68.2	376 1	THIL_ZOORA
25	30	68.2	391 1	THIL_ZOORA
26	30	68.2	393 1	RP42_THECE
27	30	68.2	480 1	GLG1_RHIME
28	30	68.2	882 1	KEU2_YEAST
29	30	68.2	912 1	ICAS_RABIT
30	30	68.2	916 1	GYRA_NEIGO
31	30	68.2	924 1	ICAS_HUMAN
32	30	68.2	1164 1	KEU1_YEAST
33	30	68.2	4499 1	DYHA_CHURE

34	29	65.9	106 1	BOLA_VIBCH	Q9KPS0 vibrio chol
35	29	65.9	130 1	RL3_PIG	Q29293 sus scrofa
36	29	65.9	149 1	ENRN_BPT7	P00641 bacterioph
37	29	65.9	229 1	TDX1_BRUMA	P48822 brugia mela
38	29	65.9	235 1	TYRF_BARRA	O44843 bartonella
39	29	65.9	268 1	TYRF_ARATH	P26587 arabidopsis
40	29	65.9	289 1	TYRF_CHICK	P25324 gallus gall
41	29	65.9	292 1	CIRG_ECO57	P58161 escherichia
42	29	65.9	292 1	CIRG_ECOLI	P77231 escherichia
43	29	65.9	306 1	RS1_SYNP6	P46228 synechococ
44	29	65.9	328 1	RS1A_SYNY3	P73530 synechocyst
45	29	65.9	334 1	Y111_HAINT	P23464 halobacteri

## ALIGNMENTS

```

RESULT 1
PRRP_RAT
ID PRRP_RAT STANDARD; PRT; 83 AA.
AC P81278;
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE Proactin-releasing peptide precursor (PRRP) (Proactin-releasing
DE hormone) [Contains: Proactin-releasing peptide PRRP31; Proactin-
DE releasing peptide PRRP20].
GN PRR.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=98268781; PubMed=9607765;
RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,
RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,
RA Kurokawa T., Nishimura O., Onda H., Fujino M.;
RT "A proactin-releasing peptide in the brain.";
RL Nature 393:272-276(1998).
RN [2]
RP TISSUE SPECIFICITY.
RX PubMed=10498338;
RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,
RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,
RA Sumino Y., Fujino M.;
RT "Tissue distribution of proactin-releasing peptide (PRRP) and its
RT receptor.";
RL Regul. Pept. 83:1-10(1999).
CC -I- FUNCTION: Stimulates proactin (PRL) release and regulates the
CC expression of proactin through its receptor GPR10. May stimulate
CC lactotrophs directly to secrete PRL.
CC -I- TISSUE SPECIFICITY: Widely expressed, with highest levels in
CC medulla oblongata and hypothalamus.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC
DR EMBL, AB015418; BAA29026.1; -
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.
FT SIGNAL 1 21
FT PEPTIDE 22 52 PROACTIN-RELEASING PEPTIDE PRRP31.
FT PEPTIDE 33 52 PROACTIN-RELEASING PEPTIDE PRRP20.
FT MOD RES 52 52 AMIDATION (G-53 PROVIDE AMIDE GROUP).
SQ SEQUENCE 83 AA; 9215 MW; DQC75A264EEBAF29 CXC64;

```

Query Match 97.7%; Score 43; DB 1; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.023;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPYGRF 8  
|||||  
Db 45 GIRPYGRF 52

## RESULT 2

PRRP\_HUMAN

ID PRRP\_HUMAN STANDARD: PRT: 87 AA.

AC P8127;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].  
DE PRR.

OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;

[1] SEQUENCE FROM N.A.

RC MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kurokawa T., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,

RT "A proactin-releasing peptide in the brain.";

RL Nature 393:272-276(1998).

RN TISSUE SPECIFICITY.

RP PubMed=10498338;

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S., Sekiguchi M., Kurokawa T., Nishimura O., Onda H.,

RT "Tissue distribution of proactin-releasing peptide (PRP) and its receptor.";

RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.  
CC TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AB015419; BAA29027.1; -  
KW MIM; 602663; -

DR MIM; 602663; -  
KW Hormone; Amidation; Signal.

FT SIGNAL 1 22 BY SIMILARITY.  
FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP1.  
FT PEPTIDE 34 53 PROACTIN-RELEASING PEPTIDE PRRP2.

FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 87 AA; 9639 MW; 229A2F3F50CF981B CRC64;

Query Match 97.7%; Score 43; DB 1; Length 87;  
Best Local Similarity 100.0%; Pred. No. 0.024;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPYGRF 8  
|||||  
Db 46 GIRPYGRF 53

## RESULT 3

PRRP\_BOVIN

ID PRRP\_BOVIN STANDARD: PRT: 98 AA.

AC P81264;  
DT 30-MAY-2000 (Rel. 39, Created)  
DT 30-MAY-2000 (Rel. 39, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Proactin-releasing peptide precursor (PRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-releasing peptide PRRP2].  
DE PRR.

OS Bos taurus (Bovine).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
OC Bovidae; Bovinae; Bos.

OX NCBI\_TaxID=9913;  
[1] SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

RP TISSUE=Brain;  
RC MEDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S., Kurokawa T., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,

RT "A proactin-releasing peptide in the brain.";

RL Nature 393:272-276(1998).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the expression of prolactin through its receptor GPR10. May stimulate lactotrophs directly to secrete PRL.

CC TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL; AB015417; BAA29025.1; -  
KW Hormone; Amidation; Signal; Cleavage on pair of basic residues.

FT SIGNAL 1 22  
FT PEPTIDE 23 53 PROACTIN-RELEASING PEPTIDE PRRP1.  
FT PEPTIDE 33 53 PROACTIN-RELEASING PEPTIDE PRRP2.

FT MOD\_RES 53 53 AMIDATION (G-54 PROVIDE AMIDE GROUP).  
SQ SEQUENCE 98 AA; 10544 MW; 08AC35A13B0FA908 CRC64;

Query Match 97.7%; Score 43; DB 1; Length 98;  
Best Local Similarity 100.0%; Pred. No. 0.027;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPYGRF 8  
|||||  
Db 46 GIRPYGRF 53

## RESULT 4

YGC1\_ECOLI

ID YGC1\_ECOLI STANDARD: PRT: 224 AA.

AC Q46898;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical protein ygc1 precursor.

GN YGC1 OR B2757.  
OS Escherichia coli.

OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
OC Escherichia.

OX NCBI\_TaxID=562;  
[1] SEQUENCE FROM N.A.

RP STRAIN=K12 / MG1655;



RX MEDLINE-97426617; PubMed-9278503;  
RA Blatter F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
RA Riley M., Colado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
RA Mau B., Shao Y.;  
RT "The complete genome sequence of *Escherichia coli* K-12.";  
RL Science 277:1453-1474(1997).  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: U29579; AAA69267.1; ALT\_INIT.  
DR EMBL: AE000359; AAC75799.1; ALT\_INIT.  
DR Ecogene: E613116; ygc1.  
KM Hypothetical protein; Signal; Complete proteome.  
FT SIGNAL 1 17  
FT CHAIN 18 224  
SQ SEQUENCE 224 AA; 25209 MW; 1C42CC009B317D68 CRC64;

Query Match 77.3%; Score 34; DB 1; Length 224;  
Best Local Similarity 75.0%; Pred. No. 5.2;  
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

OY 1 GIRPVGR 8  
1 1 1 1 1 1 1 1  
DB 23 GTRPGGR 30

RESULT 5  
GYRA\_AERSA STANDARD; PRT; 922 AA.  
ID GYRA\_AERSA  
AC P48369;  
DT 01-FEB-1996 (Rel. 33, Created)  
DT 01-FEB-1996 (Rel. 33, Last sequence update)  
DT 01-NOV-1997 (Rel. 35, Last annotation update)  
DE DNA gyrase subunit A (EC 5.99.1.3).  
GN GYRA.  
OS Aeromonas salmonicida.  
OC Bacteria; Proteobacteria; gamma subdivision; Aeromonadaceae;  
OC Aeromonas.  
OX NCBI\_TaxID=645;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-2148/89;  
RA Opegaard H.;  
RL Submitted (OCT-1995) to the EMBL/Genbank/DBJ databases.  
RN [2]  
RP SEQUENCE OF 33-179 FROM N.A.  
RC STRAIN-ATCC 14174;  
RX MEDLINE-95142596; PubMed-7840589;  
RA Opegaard H., Sorum H.;  
RT "gyrA mutations in quinolone-resistant isolates of the fish pathogen  
RT *Aeromonas salmonicida*.";  
RL Antimicrob. Agents Chemother. 38:2460-2464(1994).  
RL Antimicrob. Agents Chemother. 38:2460-2464(1994).  
CC -----  
CC -1- FUNCTION: DNA GYRASE NEGATIVELY SUPERCOILS CLOSED CIRCULAR DOUBLE-  
CC STRANDED DNA IN AN ATP-DEPENDENT MANNER AND ALSO CATALYZES THE  
CC INTERCONVERSION OF OTHER TOPOLOGICAL ISOMERS OF DOUBLE-STRANDED  
CC DNA RINGS, INCLUDING CATENANES AND KNOTTED RINGS.  
CC -1- CATALYTIC ACTIVITY: ATP-dependent breakage, passage and rejoining  
CC of double-stranded DNA.  
CC -1- SUBUNIT: MADE UP OF TWO CHAINS. THE A CHAIN IS RESPONSIBLE FOR DNA  
CC BREAKAGE AND REJOINING. THE B CHAIN CATALYZES ATP HYDROLYSIS. THE  
CC ENZYME FORMS AN A2B2 TETRAMER.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL: L47978; ABA41037.1; -;  
DR EMBL: L42453; AAB87239.1; -;  
DR HSSP: P09097; IAB4.  
DR InterPro: IPR002205; DNA\_topoisomV.  
DR Pfam: PF00521; DNA\_topoisomV; 1.  
DR SMART: SM00434; TOP4c; 1.  
DR Isomerase; Topoisomerase; DNA-binding.  
FT ACT\_SITE 122 122  
SQ SEQUENCE 922 AA; 10133 MW; 889495DC4217077 CRC64;

Query Match 75.0%; Score 33; DB 1; Length 922;  
Best Local Similarity 71.4%; Pred. No. 36;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
1 1 1 1 1 1 1 1  
DB 759 GYRPMGR 765

RESULT 6  
AR12\_HUMAN STANDARD; PRT; 330 AA.  
ID AR12\_HUMAN  
AC O43488; O75749;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Aflatoxin B1 aldehyde reductase 1 (EC 1.-.-.-) (AFB1-AR 1)  
DE (Aldoketoreductase 7).  
GN AKR7A2 OR AFAR OR AKR7.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE-98244807; PubMed-9576847;  
RA Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;  
RT "Molecular cloning, expression and catalytic activity of a human AKR7  
RT member of the aldo-keto reductase superfamily: evidence that the  
RT major 2-carboxybenzaldehyde reductase from human liver is a homologue  
RT of rat aflatoxin B1-aldehyde reductase.";  
RL Biochem. J. 332:21-34(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE-99040634; PubMed-9823300;  
RA Pram C., Saveiyeva L., Petri P., Schwab M.;  
RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-  
RT 1p36.1 in a region frequently altered in human tumor cells.";  
RL Cancer Res. 58:5014-5018(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (FEB-2000) to the EMBL/Genbank/DBJ databases.  
CC -----  
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING  
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOL. COULD BE  
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC  
CC EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE  
CC REDUCTASE.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF026947; AAC52104.1; -  
DR EMBL: Y16675; CAAT5347.1; -  
DR EMBL: AL035413; CAB72321.1; -  
DR MIM: 603418; -  
DR InterPro: IPR001395; Aldo\_ket\_red.  
DR Pfam: PF00248; aldo\_ket\_red; 1.  
KM Oxidoreductase.  
FT ACT\_SITE 112 112 HYDROGEN-BOND DONOR (PROBABLE).  
FT CONFLICT 113 113 A -> T (IN REF. 1).  
SQ SEQUENCE 330 AA; 36618 MW; 3BBFB7ED0CA4D54 CRC64;  
  
Query Match 72.7% Score 32; DB 1; Length 330;  
Best Local Similarity 75.0%; Pred. NO. 20;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GIRPVGR 8  
I:|||||  
Db 215 GKQPVGR 222  
  
RESULT 7  
AR73\_HUMAN STANDARD; PRT; 331 AA.  
AC 095154; O9NUC3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Alkaline B1 aldehyde reductase 2 (EC 1.1.1.1) (AFB1-AR 2).  
GN AR7A3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.  
RC TISSUE=Liver;  
RX MEDLINE=99315412; PubMed=10383892;  
RA Knight L.P., Primiano T., Groopman J.D., Kensler T.W., Sutter T.R.;  
RT cDNA cloning, expression and activity of a second human alatoxin  
RT B1-metabolizing member of the aldo-keto reductase superfamily,  
RT AR7A3.";  
RL Carcinogenesis 20:1215-1223(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
CC -!- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING  
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOLCHOL. COULD BE  
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC  
CC EFFECTS OF AFLATOXIN B1.  
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -!- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF040639; AAD02195.1; -  
DR EMBL: AL035413; CAB72322.1; -  
DR InterPro: IPR001395; Aldo\_ket\_red.  
DR Pfam: PF00248; aldo\_ket\_red; 1.  
DR PRINTS: PR00069; ALDKERDRTASE.  
KM Oxidoreductase.  
FT ACT\_SITE 113 113 HYDROGEN-BOND DONOR (PROBABLE).  
SQ SEQUENCE 113 113

FT CONFLICT 51 51 E -> D (IN REF. 1).  
FT CONFLICT 138 138 V -> M (IN REF. 1).  
FT CONFLICT 201 201 A -> ADQSEGGCSFMTLGPADCFPS (IN REF.  
FT CONFLICT 215 215 2).  
FT CONFLICT 205 215 N -> D (IN REF. 1).  
FT CONFLICT 323 323 T -> A (IN REF. 1).  
SQ SEQUENCE 331 AA; 37206 MW; B9C32C33C7102AB3 CRC64;  
  
Query Match 72.7% Score 32; DB 1; Length 331;  
Best Local Similarity 75.0%; Pred. NO. 21;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GIRPVGR 8  
I:|||||  
Db 216 GKQPVGR 223  
  
RESULT 8  
YD05\_SCHPO STANDARD; PRT; 207 AA.  
ID YD05\_SCHPO  
AC 014198;  
DT 15-JUL-1998 (Rel. 36, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Hypothetical 24.0 kDa protein C5D6.05 in chromosome I.  
GN SPAC5D6.05.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Skelton J., Churcher C.M., Wood V., Barrell B.G., Rajandream M.A.;  
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: Z98056; CAB10853.2; -  
KM Hypothetical protein.  
SQ SEQUENCE 207 AA; 24027 MW; A3A3B518E3D85A62 CRC64;  
  
Query Match 70.5% Score 31; DB 1; Length 207;  
Best Local Similarity 85.7%; Pred. NO. 21;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
QY 1 GIRPVGR 7  
I:|||||  
Db 95 GIRPVNR 101  
  
RESULT 9  
PYRF\_HAEIN STANDARD; PRT; 230 AA.  
ID PYRF\_HAEIN  
AC P43812;  
DT 01-NOV-1995 (Rel. 32, Created)  
DT 01-NOV-1995 (Rel. 32, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Oxidative 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)  
DE (OMPDecase) (OMPDecase).  
GN PYRF OR H11224.  
OS Haemophilus influenzae.  
OC Bacterii; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Haemophilus.

OX NCBI\_TaxID=727;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-RD / KM20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kiehlmann A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McElveney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shiley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Uterback T.R., Hanna M.C., Nguyen D.T., Saudke D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Georgiades N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus  
RT influenzae Rd.";  
RL Science 269:496-512(1995).  
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate - UMP + CO(2).  
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; U32802; AAC22877.1; -.  
DR TIGR; H11224; -.  
DR InterPro; IPR001754; OMPdecase.  
DR Pfam; PF00215; OMPdecase; 1.  
DR PROSITE; PS00156; OMPDECASE; 1.  
KW Lyase: Decarboxylase; Pyrimidine biosynthesis: Complete proteome.  
FT ACT\_SITE 61 61 PROTON DONOR (BY SIMILARITY).  
SQ SEQUENCE 230 AA; 25192 MW; 14EBE825DBC767D9 CRC64;  
  
Query Match 70.5%; Score 31; DB 1; Length 230;  
Best Local Similarity 83.3%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GIRPYG 6  
Db 178 GIRPYG 183  
||||:|  
178 GIRPYG 183  
  
RESULT 10  
PYRF\_PASMU STANDARD; PRT; 233 AA.  
AC 09CMM1;  
DT 01-MAR-2002 (Rel. 41, Created)  
DT 01-MAR-2002 (Rel. 41, Last sequence update)  
DT 01-MAR-2002 (Rel. 41, Last annotation update)  
DE Orotidine 5'-phosphate decarboxylase (EC 4.1.1.23) (OMP decarboxylase)  
DE (OMPdecase) (OMPdecase).  
GN PYRF OR PM0797.  
OS Pasteurella multocida.  
OC Bacteria; Proteobacteria; gamma subdivision; Pasteurellaceae;  
OC Pasteurella.  
OX NCBI\_TaxID=747;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-Pm70;  
RX MEDLINE=21145866; PubMed=11248100;  
RA May B.-J., Zhang Q., Li L.L., Paustian M.L., Whittam T.S., Kapur V.;  
RT "Complete genomic sequence of Pasteurella multocida Pm70.";  
RL Proc. Natl. Acad. Sci. U.S.A. 98:3460-3465(2001).  
CC -1- CATALYTIC ACTIVITY: Orotidine 5'-phosphate - UMP + CO(2).  
CC -1- PATHWAY: Pyrimidine biosynthesis; sixth (last) step.  
CC -1- SUBUNIT: Homodimer (By similarity).  
CC -----

CC -1- SIMILARITY: BELONGS TO THE OMP DECARBOXYLASE FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; AE006118; AAK02881.1; -.  
DR InterPro; IPR001754; OMPdecase.  
DR Pfam; PF00215; OMPdecase; 1.  
DR PROSITE; PS00156; OMPDECASE; 1.  
KW Lyase: Decarboxylase; Pyrimidine biosynthesis: Complete proteome.  
FT ACT\_SITE 61 61 PROTON DONOR (BY SIMILARITY).  
SQ SEQUENCE 233 AA; 25501 MW; 6CD84E5931A60628 CRC64;  
  
Query Match 70.5%; Score 31; DB 1; Length 233;  
Best Local Similarity 83.3%; Pred. No. 23;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
  
QY 1 GIRPYG 6  
Db 178 GIRPYG 183  
||||:|  
178 GIRPYG 183  
  
RESULT 11  
LINC\_PSEPA STANDARD; PRT; 250 AA.  
AC P50197;  
DT 01-OCT-1996 (Rel. 34, Created)  
DT 01-OCT-1996 (Rel. 34, Last sequence update)  
DT 01-OCT-1996 (Rel. 34, Last annotation update)  
DE 2,5-dichloro-2,5-cyclohexadiene-1,4-diol dehydrogenase (EC 1.1.-.-)  
DE (2,5-DDOL dehydrogenase).  
GN LINC.  
OS Pseudomonas paucimobillis (Sphingomonas paucimobillis).  
OC Bacteria; Proteobacteria; alpha subdivision; Sphingomonadaceae;  
OC Sphingomonas.  
OX NCBI\_TaxID=13689;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-UT26;  
RX MEDLINE=94252977; PubMed=7515041;  
RA Nagata Y., Ohtomo R., Miyauchi K., Fukuda M., Yano K., Takagi M.;  
RT "Cloning and sequencing of a 2,5-dichloro-2,5-cyclohexadiene-1,4-diol  
RT dehydrogenase gene involved in the degradation of gamma-  
RT hexachlorocyclohexane in Pseudomonas paucimobillis.";  
RL J. Bacteriol. 176:3117-3125(1994).  
DE -1- FUNCTION: DEGRADATION OF 2,5-DICHLORO-2,5-CYCLOHEXADIENE-1,4-DIOL  
DE (2,5-DDOL) INTO 2,5-DICHLOROHEXANOINONE (2,5-DCHO).  
DE -1- PATHWAY: DEGRADATION OF GAMMA-HEXACHLOROCYCLOHEXANE.  
DE (SDR) FAMILY.  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
CC -----  
DR EMBL; D14595; BAA03444.1; -.  
DR HSSP; P19992; IHDC.  
DR InterPro; IPR002198; ADH\_short.  
DR Pfam; PF00106; adh\_short; 1.  
DR PRINTS; PR00080; SDRFAMILY.  
DR PROSITE; PS00061; ADH\_SHORT; 1.  
KW Aromatic hydrocarbons catabolism; Oxidoreductase; NAD.  
FT NP\_BIND 9 34 NAD (BY SIMILARITY).  
CC -----

FT ACT\_SITE 154 154 BY SIMILARITY.  
 SQ SEQUENCE 250 AA; 25644 MW; FFC1CAEB47DF789D CRC64;

Query Match  
 Best Local Similarity 70.5%; Score 31; DB 1; Length 250;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 2 IRPVGRF 8  
 1 1 1 1 1  
 DB 209 IHPVGRF 215

RESULT 12  
 CSD\_THEME STANDARD; PRT; 413 AA.  
 ID CSD\_THEME 09X191;

DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Probable cysteine desulfurase (EC 4.4.1.-).  
 GN CSD OR TM1371.  
 OS Thermotoga maritima.  
 OC Bacteria; Thermotogales; Thermotoga.  
 OX NCBI\_TaxID=2336.

RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
 McDonald L., Uetebach T.R., Malek J.A., Linner K.D., Garrett M.W.,  
 Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
 Heideberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
 Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
 RT Evidence for lateral gene transfer between Archaea and Bacteria from  
 RT genome sequence of Thermotoga maritima.";

RL Nature 399:323-329(1999).  
 CC -1- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR AND SELENIUM  
 CC ATOMS FROM L-CYSTEINE, L-CYSTEINE, L-SELENOCYSTEINE, AND L-  
 CC SELENOCYSTEINE TO PRODUCE L-ALANINE (BY SIMILARITY).  
 CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
 CC AMINOTRANSFERASES. CSD SUBFAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL; AE001791; AAD36460.1; -;  
 DR HSSP; P77444; ICON.  
 DR TIGR; TM1371; -;  
 DR InterPro; IPR000192; AminoTransf\_class\_V.  
 DR Pfam; PF00266; aminoTransf\_5; 1.  
 DR PROSITE; PS00595; AA\_TRANSF\_CLASS\_5; FALSE\_NEG.  
 DR Lyase; Pyridoxal phosphate; Complete proteome.  
 FT BINDING 223 223 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
 SQ SEQUENCE 413 AA; 46748 MW; 092866323FF984F0 CRC64;

Query Match  
 Best Local Similarity 70.5%; Score 31; DB 1; Length 413;  
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 IRPVGRF 8  
 1 1 1 1 1  
 DB 142 IRPSGRF 148

RESULT 13  
 ID SYE\_BACHD STANDARD; PRT; 485 AA.  
 AC 09KGF6;

DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glutamyl-tRNA synthetase (EC 6.1.1.17) (Glutamate--tRNA ligase)  
 DE (GLURS).  
 GN GLTX OR BH0109.  
 OS Bacillus halodurans.  
 OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
 OC Bacillus/Staphylococcus group; Bacillus.  
 OX NCBI\_TaxID=6665;  
 OX [1]

RA STRAIN-C-125 / JCM 9153;  
 RX MEDLINE-20512582; PubMed-11058132;  
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
 Fujii F., Hirama C., Nakamura Y., Ogasawara N., Kuhara S.,  
 Horikoshi K.;  
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";  
 RL Nucleic Acids Res. 28:4317-4331(2000).  
 CC -1- CATALYTIC ACTIVITY: ATP + L-glutamate + tRNA(Glu) -> AMP +  
 CC diphosphate + L-glutamyl-tRNA(Glu).  
 CC -1- SUBUNIT: MONOMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO CLASS-I AMINOACYL-TRNA SYNTHETASE FAMILY.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

CC EMBL; AP001507; BAB03828.1; -;  
 DR InterPro; IPR000924; tRNA-synt\_1c.  
 DR InterPro; IPR001412; tRNA-synt\_1.  
 DR Pfam; PF00749; tRNA-synt\_1c; 1.  
 DR PRINTS; PR00987; TRNASYNTGCU.  
 DR PROSITE; PS00178; AA-TRNA\_LIGASE\_I; 1.  
 DR Aminoacyl-tRNA synthetase; Protein biosynthesis; Ligase; ATP-binding;  
 KW Complete proteome.  
 FT SITE 11 21 "HIGH" REGION.  
 FT SITE 252 256 "KMSK" REGION.  
 FT BINDING 255 255 ATP (BY SIMILARITY).  
 SQ SEQUENCE 485 AA; 54785 MW; 7D34A862918F57B6 CRC64;

Query Match  
 Best Local Similarity 70.5%; Score 31; DB 1; Length 485;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPVGRF 8  
 1 1 1 1 1  
 DB 150 GIRPVGRF 157

RESULT 14  
 DHGL\_DROME STANDARD; PRT; 625 AA.  
 ID DHGL\_DROME

AC P18173; 09V187;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glucose dehydrogenase [acceptor] precursor (EC 1.1.99.10) [Contains:  
 DE Glucose dehydrogenase [acceptor] short protein].  
 GN GMD OR CG1152.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;

OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 RX NCBI\_TaxID=7227;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90205602; PubMed=2108306;  
 RA Krasney P.A., Carr C.M., Cavenier D.R.;  
 RT "Evolution of the glucose dehydrogenase gene in Drosophila.";  
 RL Mol. Biol. Evol. 7:155-177(1990).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-BERKELEY;  
 RX MEDLINE=20196006; PubMed=10731132;  
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
 RA Brandon R.C., Rogers Y.H.C., Blazer R.G., Champagne M., Pfeiffer B.D.,  
 RA Wan K.H., Doyle C., Barker E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
 RA Abill J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
 RA Beeson K.Y., Bens P.V., Bertman B.P., Bhandari D., Bolshakov S.,  
 RA Borkova D., Botchan M.R., Bouck J., Brocktein P., Brotler P.,  
 RA Burris K.C., Busan D.A., Butler H., Cadieu E., Center A., Chandra I.,  
 RA Cherry J.M., Cawley S., Dahlke C., Davaport L.B., Davies P.,  
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
 RA Durkin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
 RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
 RA Harris N.L., Harvey D., Helman T.J., Hernandez J.R., Houck J.,  
 RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,  
 RA Jajuli M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
 RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
 RA Merulov G., Milshina N.V., Modarri C., Morris J., Moshrefi A.,  
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
 RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
 RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
 RA Svitskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
 RA Wang Z.-Y., Wasserman D.A., Welnslock G.M., Weissbach J.,  
 RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
 RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,  
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;  
 RT "The genome sequence of Drosophila melanogaster.";  
 RL Science 287:2185-2195(2000).  
 RN [3]  
 RP SEQUENCE OF 1-96 FROM N.A.  
 RX MEDLINE=89065557; PubMed=3143620;  
 RA Whetten R., Organ E., Krasney P., Cox-Foster D., Cavenier D.R.;  
 RT "Molecular structure and transformation of the glucose dehydrogenase  
 gene in Drosophila melanogaster.";  
 RL Genetics 120:475-484(1988).  
 RN [4]  
 RP SIMILARITY TO YEAST ALCOHOL OXIDASE.  
 RX MEDLINE=91163320; PubMed=2002763;  
 RA Cavenier D.R., Krasney P.;  
 RT "Drosophila glucose dehydrogenase and yeast alcohol oxidase are  
 homologous and share N-terminal homology with other flavoenzymes.";  
 RL Mol. Biol. Evol. 8:144-150(1991).  
 RN [5]  
 RP SELENOCYSTEINE.  
 RA Perlicky S., Merritt K., Cavenier D.;  
 RT "Incorporation of selenocysteine at a UGA codon of Gld.";  
 RL (In) Annu. Dros. Res. Conf. 39:414C(1998).  
 CC -1- FUNCTION: ESSENTIAL FOR CUTICULAR MODIFICATION DURING DEVELOPMENT.  
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +  
 reduced acceptor.

CC -1- CORACTOR: FAD.  
 CC -1- SUBCELLULAR LOCATION: SECRETED AS PART OF THE SEMINAL FLUID  
 CC TRANSFERRED TO FEMALES.  
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.  
 CC -1- CAUTION: REF.5 BELIEVES RESIDUE 613 IS A SELENOCYSTEINE.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb.ch/announce/>  
 CC or send an email to [license@isb.sib.ch](mailto:license@isb.sib.ch)).  
 CC -----  
 CC EMBL: M29298; AAA28571.1; ALT\_SEQ.  
 CC EMBL: AE003672; AAF54038.1; ALT\_SEQ.  
 CC EMBL: M13582; GAA31918.1; -.  
 CC PIR: A39019; A39019.  
 CC PIR: S06628; S06628.  
 CC PIR: FBgn0001112; Gld.  
 CC FlyBase: FBgn0001112; Gld.  
 CC InterPro: IPR000172; GMC\_oxred.  
 CC Pfam: PF00732; GMC\_oxred.  
 CC PROSITE: PS00623; GMC\_OXRED\_1; 1.  
 CC PROSITE: PS00624; GMC\_OXRED\_2; 1.  
 CC OXfold: OXfold; GMC\_oxred.  
 CC SIGNAL 1 42  
 CC CHAIN 43 625 GLUCOSE DEHYDROGENASE [ACCEPTOR].  
 CC CHAIN 43 612 GLUCOSE DEHYDROGENASE [ACCEPTOR] SHORT  
 CC FT NP\_BIND 66 95 PROTEIN.  
 CC FT ACT\_SITE 549 549 FAD (ADP PART) (PROBABLE).  
 CC FT SE\_CYS 613 613 POTENTIAL.  
 CC FT CONFLICT 484 484 PROBABLE.  
 CC FT SEQUENCE 625 AA; 68387 MW; 02664FC3B820D6FC CRC64;  
 SO  
 Query Match 70.5%; Score 31; DB 1; Length 625;  
 Best Local Similarity 62.5%; Pred. No. 64;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GTRPVGRF 8  
 DB 207 GLUPVGRF 214  
 ID DHGL\_DROPS STANDARD; PRT; 625 AA.  
 DHGL\_DROPS  
 AC P18172;  
 DT 01-NOV-1990 (Rel. 16, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Glucose dehydrogenase [acceptor] precursor (EC 1.1.99.10) [contains:  
 DE Glucose dehydrogenase [acceptor] short protein].  
 GN GLD.  
 OS Drosophila pseudoobscura (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidea; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7237;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=90205602; PubMed=2108306;  
 RA Krasney P.A., Carr C.M., Cavenier D.R.;  
 RT "Evolution of the glucose dehydrogenase gene in Drosophila.";  
 RL Mol. Biol. Evol. 7:155-177(1990).  
 RN [2]  
 RP SELENOCYSTEINE.  
 RA Perlicky S., Merritt K., Cavenier D.;  
 RT "Incorporation of selenocysteine at a UGA codon of Gld.";  
 RL (In) Annu. Dros. Res. Conf. 39:414C(1998).  
 CC -1- CATALYTIC ACTIVITY: D-glucose + acceptor = D-glucono-1,5-lactone +  
 reduced acceptor.

CC -1- COPACTOR: FAD.  
 CC -1- SUBCELLULAR LOCATION: SECRETED AS PART OF THE SEMINAL FLUID  
 CC TRANSFERRED TO FEMALES.  
 CC -1- SIMILARITY: BELONGS TO THE GMC OXIDOREDUCTASES FAMILY.  
 CC -1- CAUTION: REF.2 BELIEVES RESIDUE 613 IS A SELENOCYSTEINE.  
 -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 -----  
 CC  
 DR EMBL; M29299; AAA28572.1; ALT\_SEQ.  
 DR PIR; B39019; B39019.  
 DR FlyBase; FBgn0012699; Dpse\Gld.  
 DR InterPro; IPR000172; GMC\_oxred.  
 DR Pfam; PF00732; GMC\_oxred; 1.  
 DR PROSITE; PS00623; GMC\_OXRED\_1; 1.  
 DR PROSITE; PS00624; GMC\_OXRED\_2; 1.  
 KW Oxidoreductase; Flavoprotein; FAD; Signal; Selenium; Selenocysteine.  
 FT SIGNAL 1 42  
 FT CHAIN 43 625 BY SIMILARITY.  
 FT CHAIN 43 612 GLUCOSE DEHYDROGENASE [ACCEPTOR].  
 FT NP\_BIND 66 95 PROTEIN.  
 FT ACT\_SITE 549 549 FAD (ADP PART) (PROBABLE).  
 FT SE\_CYS 613 613 POTENTIAL.  
 FT SEQUENCE 625 AA; 68557 MW; C5523E97CFD73B9 CRC64;  
 SO PROBABLE.

Query Match 70.5%; Score 31; DB 1; Length 625;  
 Best Local Similarity 62.5%; Pred. No. 64;  
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPYGRF 8  
 DB 207 GLLPVGRF 214

Search completed: September 13, 2002, 09:30:49  
 Job time: 1139 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:24 ; Search time 311.85 Seconds  
(without alignments)  
4.993 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_13\_21

Perfect score: 44

Sequence: 1 GIRPVGREF 9

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP mnc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*  
15: SP\_rvirus: \*  
16: SP\_bacteriophage: \*  
17: SP\_archaeap: \*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	43	97.7	420	2	093LZ7
2	41	93.2	117	13	09W624
3	36	81.8	426	17	0979S0
4	36	81.8	674	16	050431
5	35	79.5	504	16	053895
6	35	79.5	519	16	0925G7
7	35	79.5	1056	16	09PA00
8	35	79.5	1172	5	09VU79
9	34	77.3	529	3	09C1M8
10	33	75.0	139	5	09N6S1
11	33	75.0	353	16	033194
12	33	75.0	366	16	066950
13	33	75.0	445	2	050531
14	33	75.0	511	3	059714
15	33	75.0	894	16	09KSJ8
16	32	72.7	183	16	092KK8

17	32	72.7	285	2	09X4W7	09x4w7 pseudomonas
18	32	72.7	292	16	092VM1	092vm1 rhizobium m
19	32	72.7	299	16	P95060	P95060 mycobacteri
20	32	72.7	309	2	086546	086546 streptomyc
21	32	72.7	315	11	09DCM1	09dcm1 mus musculi
22	32	72.7	333	2	09RJT0	09rjt0 streptomyc
23	32	72.7	356	16	0986C9	0986c9 rhizobium l
24	32	72.7	366	11	09D157	09d157 mus musculi
25	32	72.7	374	16	053319	053319 mycobacteri
26	32	72.7	409	2	09KK77	09kk77 brevibacter
27	32	72.7	435	17	09VE81	09ve81 aeropyrum p
28	32	72.7	439	10	09FKJ4	09fkj4 arabidopsis
29	32	72.7	439	10	0940M3	0940m3 arabidopsis
30	32	72.7	492	16	010884	010884 mycobacteri
31	32	72.7	704	10	09FF49	09ff49 arabidopsis
32	32	72.7	869	16	098B21	098b21 rhizobium l
33	32	72.7	1057	13	P79999	P79999 oryzae lat
34	32	72.7	1059	5	019370	019370 caenorhabdi
35	31	70.5	125	16	09A6F2	09a6f2 caulobacter
36	31	70.5	129	16	09KUJ6	09kuj6 vibrio chol
37	31	70.5	148	12	09DWG5	09dwg5 rat cytomeg
38	31	70.5	163	2	045285	045285 bacillus l1
39	31	70.5	172	3	096VP9	096vp9 schizosacch
40	31	70.5	233	16	09CMM1	09cmm1 pasteurella
41	31	70.5	258	2	005664	005664 listeria mo
42	31	70.5	286	16	P95150	P95150 mycobacteri
43	31	70.5	308	16	0928H6	0928h6 listeria in
44	31	70.5	319	2	088067	088067 streptomyc
45	31	70.5	352	2	09EW95	09ew95 streptomyc

## ALIGNMENTS

RESULT 1  
ID 093LZ7 PRELIMINARY; PRT; 420 AA.  
AC 093LZ7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CCM3239;  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CCM3239.";  
RL Submitted (Mar-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AY033994; AAK61719.1; -  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE8C2DEA CRC64;

Query Match 97.7%; Score 43; DB 2; Length 420;  
Best Local Similarity 100.0%; Pred. No. 1;  
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 GIRPVGREF 8  
Db 48 GIRPVGREF 55  
RESULT 2  
ID 09W624 PRELIMINARY; PRT; 117 AA.  
AC 09W624;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE C-RE AMIDE.  
 OS Carassius auratus (Goldfish).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Carassius.  
 NC NCBI\_taxid=7937;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE-BRAIN.  
 RA Satake H., Minakata H., Fujimoto M.;  
 RT "Carassius Karamida (C-RE amide).";  
 RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AB020024; BAA7662.1; -;  
 SO SEQUENCE 117 AA; 12879 MW; D5DC4CB23038C3B0 CRC64;

Query Match 93.2%; Score 41; DB 13; Length 117;  
 Best Local Similarity 75.0%; Pred. No. 0.68;  
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGR 8  
 I:|:|:|:|  
 Db 68 GVRIGRF 75

RESULT 3  
 Q97950 PRELIMINARY; PRT; 426 AA.  
 AC 097950;  
 DT 01-OCT-2001 (TReMBLrel. 18, Created)  
 DT 01-OCT-2001 (TReMBLrel. 18, Last sequence update)  
 DE 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DT TVG1120468 PROTEIN.  
 GN TVG1120468.  
 OS Thermoplasma volcanium.  
 OC Archaea; Euryarchaeota; Thermoplasmatales; Thermoplasmaceae;  
 OC Thermoplasm.  
 NC NCBI\_taxid=50339;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-GSSI / DSM 4299 / JCM 9571;  
 RX MEDLINE-20570466; PubMed-11121031;  
 RA Kawashima T., Amano N., Koike H., Makino S.-I., Higuchi S.,  
 RA Nunoshima T., Yamamoto Y., Aramaki H., Kanehori K., Kawamoto T.,  
 RT "Archaeal adaptation to higher temperatures revealed by genomic  
 RT sequence of Thermoplasma volcanium.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 97:14257-14262(2000).  
 DR EMBL: AP000995; BAB6032.1; -;  
 DR InterPro: IPR002881; DUF58.  
 DR InterPro: IPR002035; VFMA.  
 DR Pfam: PF01882; DUF58; 1.  
 DR SMART: SM00327; VMA; 1.  
 KW Complete proteome.  
 SO SEQUENCE 426 AA; 49092 MW; 63967CF711116F4A CRC64;

Query Match 81.8%; Score 36; DB 17; Length 426;  
 Best Local Similarity 75.0%; Pred. No. 29;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GIRPVGR 8  
 I:|:|:|:|  
 Db 304 GIRPVGR 311

RESULT 4  
 Q50431 PRELIMINARY; PRT; 674 AA.  
 AC 050431;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)

DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
 DE 2,4-DIENOYL-COA REDUCTASE.  
 GN PADH OR RV1175C OR MTW005.11C.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NC NCBI\_taxid=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE-98295987; PubMed-9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 DR EMBL: AL010186; CAA15852.1; -;  
 DR Tuberculist; RV1175C;  
 DR InterPro: IPR001064; Crystallin.  
 DR InterPro: IPR001327; FAD\_pyr\_redox.  
 DR InterPro: IPR002025; NAD\_binding.  
 DR InterPro: IPR001155; Oxidored\_FMN.  
 DR Pfam: PF00724; oxidored\_FMN; 1.  
 DR Pfam: PF00070; pyr\_redox; 1.  
 DR PROSITE: PS00225; CRYSTALLIN\_BETAGAMMA; UNKNOWN\_1.  
 KW Complete proteome.  
 SO SEQUENCE 674 AA; 72901 MW; 8F478F74ED3E0A5 CRC64;

Query Match 81.8%; Score 36; DB 16; Length 674;  
 Best Local Similarity 85.7%; Pred. No. 48;  
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GIRPVGR 7  
 I:|:|:|:|  
 Db 492 GVRVGR 498

RESULT 5  
 Q53895 PRELIMINARY; PRT; 504 AA.  
 ID 053895;  
 DT 01-JUN-1998 (TReMBLrel. 06, Created)  
 DT 01-JUN-1998 (TReMBLrel. 06, Last sequence update)  
 DE PUTATIVE TWO-COMPONENT SENSOR.  
 GN RV0982 OR MTW044.10.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
 NC NCBI\_taxid=1773;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RX MEDLINE-98295987; PubMed-9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrell B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence.";  
 RL Nature 393:537-544(1998).  
 CC -I- SIMILARITY: TO PROKARYOTE SENSOR TRANSDUCTION PROTEINS.



DR EMBL: AL021999; CAA17581.1; -  
DR HSSP: P02933; 1BXD.  
DR TubercuList: Rv0982; -  
DR InterPro: IPR003660; HAMF.  
DR InterPro: IPR003594; HATPase\_C.  
DR InterPro: IPR003661; His\_KinA.  
DR InterPro: IPR004359; His\_Kin\_Sig.  
DR Pfam: PF00672; HAMF; 1.  
DR Pfam: PF02518; HATPase\_C; 1.  
DR Pfam: PF00512; signal; 1.  
DR SMART: SM00304; HAMF; 1.  
DR SMART: SM00387; HATPase\_C; 1.  
DR SMART: SM00388; HisKA; 1.  
DR Complete proteome; Kinase; Phosphorylation; Sensory transduction;  
Transferrase.  
SQ SEQUENCE 504 AA; 54407 MW; 26792A71AC432232 CRC64;

Query Match 79.5%; Score 35; DB 16; Length 504;  
Best Local Similarity 85.7%; Pred. No. 57;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPYGR 7  
|:|||||  
DB 188 GIRPYGR 194

RESULT 6  
0925G7 PRELIMINARY; PRT; 519 AA.  
ID 0925G7;  
AC 0925G7;  
DT 01-MAY-1999 (TReMBLrel. 10, Created)  
DT 01-MAY-1999 (TReMBLrel. 10, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE PUTATIVE TWO-COMPONENT SYSTEM SENSOR KINASE.  
CN MLO175 OR MLCB373.27.  
OS Mycobacterium leprae.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Corynebacterineae; Mycobacteriaceae; Mycobacterium.  
OX NCBI\_TaxID=1769;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Harris D., Taylor K.;  
RL Submitted (FEB-1999) to the EMBL/GenBank/DDBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-TN:  
RX MEDLINE-21128732; PubMed-11234002;  
RA Cole S.T., Elgineler K., Parkhill J., James K.D., Thomson N.R.,  
RA Wheeler P.R., Honore N., Garnier T., Churcher C., Harris D.,  
RA Mungall K., Basham D., Brown D., Chillingworth T., Connor R.,  
RA Davies R.M., Devlin K., Duthoy S., Feltham T., Fraser A., Hamlin N.,  
RA Holroyd S., Hornby T., Jagers K., Lacroix C., Maclean J., Moule S.,  
RA Murphy L., Oliver K., Quail M.A., Rajandream M.A., Rutherford K.M.,  
RA Rutter S., Steeger K., Simon S., Simmonds M., Skelton J., Squares R.,  
RA Squares S., Stevens K., Taylor K., Whitehead S., Woodward J.R.,  
RA Barrett B.G.;  
RT "Massive gene decay in the leprosy bacillus."  
RL Nature 409:1007-1011(2001).  
CC -1- SIMILARITY: TO PROKAROTIC SENSOR TRANSDUCTION PROTEINS.  
DR EMBL: AL035500; CAB3689.1; -  
DR EMBL: AL583917; CAC29683.1; -  
DR HSSP: P02933; 1BXD.  
DR Leprona: ML0175; -  
DR InterPro: IPR004358; BCTRLSENSOR.  
DR InterPro: IPR003660; HAMF.  
DR InterPro: IPR003594; HATPase\_C.  
DR InterPro: IPR003661; His\_KinA.  
DR Pfam: PF00672; HAMF; 1.  
DR Pfam: PF02518; HATPase\_C; 1.  
DR Pfam: PF00512; signal; 1.  
DR PRINTS: PR00344; BCTRLSENSOR.

DR SMART: SM00304; HAMF; 1.  
DR SMART: SM00387; HATPase\_C; 1.  
DR SMART: SM00388; HisKA; 1.  
DR Complete proteome; Kinase; Phosphorylation; Sensory transduction;  
Transferrase  
SQ SEQUENCE 519 AA; 56470 MW; C0F2DA057364F10 CRC64;

Query Match 79.5%; Score 35; DB 16; Length 519;  
Best Local Similarity 85.7%; Pred. No. 59;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPYGR 7  
|:|||||  
DB 188 GIRPYGR 194

RESULT 7  
09PA00 PRELIMINARY; PRT; 1056 AA.  
ID 09PA00;  
AC 09PA00;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE HYPOTHETICAL PROTEIN XP2445.  
GN XP2445.  
OS Xylella fastidiosa.  
OC Bacteria; Proteobacteria; gamma subdivision; Xanthomonas group;  
OC Xylella.  
OX NCBI\_TaxID=2371;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-9A5C;  
RX MEDLINE-20365717; PubMed-10910347;  
RA Simpson A.J.G., Relnach F.C., Arruda P., Abreu F.A., Acencio M.,  
RA Alvaranga R., Alves L.M.C., Araya J.E., Bala G.S., Baptista C.S.,  
RA Barros M.H., Bonaccorsi E.D., Bordin S., Bove J.M., Bionis M.R.S.,  
RA Bueno M.R.P., Camargo A.A., Camargo L.E.A., Carraro D.M., Carrier H.,  
RA Coulaico N.B., Colombo C., Costa F.F., Costa M.C.R., Costa-Neto C.M.,  
RA Coutinho L.L., Cristofani M., Dias-Neto E., Docena C., El-Dorri H.,  
RA Facincent A.P., Ferreira A.J.S., Ferreira V.C.A., Ferro J.A.,  
RA Fraga J.S., Franca S.C., Franco M.C., Frohme M., Furian L.R.,  
RA Garnier M., Goldman G.H., Goldman M.H.S., Gomes S.L., Gruber A.,  
RA Ho P.L., Hohlseil J.D., Junqueira M.L., Kemper E.L., Kitajima J.P.,  
RA Krieger J.E., Kuramae E.E., Laigret F., Lambais M.R., Leite L.C.C.,  
RA Lemos E.G.M., Lemos M.V.F., Lopes S.A., Lopes C.R., Machado J.A.,  
RA Machado M.A., Madella A.M.B.N., Madella H.M.F., Marino C.L.,  
RA Marques M.V., Martins E.A.L., Martins E.M.F., Matsukuma A.Y.,  
RA Menck C.F.M., Miracca E.C., Miyaki C.Y., Monteiro-Vitorello C.B.,  
RA Moon D.H., Nagai M.A., Nascimento A.L.T.O., Netto L.E.S.,  
RA Nhani A.Jr., Nobrega F.G., Nunes L.R., Oliveira M.A.,  
RA de Oliveira M.C., de Oliveira R.C., Palmieri D.A., Paris A.,  
RA Pelxoto B.R., Pereira G.A.G., Pereira H.A. Jr., Pesquero J.B.,  
RA Puggio R.B., Roberto P.G., Rodrigues V., de Rosa A.J.M.,  
RA de Rosa V.E. Jr., de Sa R.G., Santelli R.V., Sawasaki H.E.,  
RA da Silva A.C.R., da Silva A.M., da Silva F.R., Silva W.A. Jr.,  
RA da Silveira J.F., Silvestri M.L.Z., Silveira W.J., de Souza A.A.,  
RA de Souza A.P., Terenzi M.F., Truffi D., Tsai S.M., Tsuhako M.H.,  
RA Vallada H., Van Sluys M.A., Verjovski-Almeida S., Vettore A.L.,  
RA Zago M.A., Zatz M., Meidanis J., Setubal J.C.;  
RT "The genome sequence of the plant pathogen xylella fastidiosa";  
RL Nature 406:151-159(2000).  
DR EMBL: AE004053; AAF85244.1; -  
DR InterPro: IPR000421; PA58.C.  
DR InterPro: IPR001092; HTH\_dlm.  
KW Hypothetical protein; Complete proteome.  
SQ SEQUENCE 1056 AA; 118471 MW; 361CF2510BDFC10 CRC64;

Query Match 79.5%; Score 35; DB 16; Length 1056;  
Best Local Similarity 85.7%; Pred. No. 1.3e+02;  
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 GIRPVGR 7  
I:|||||  
DB 1047 GLRPVGR 1053

RESULT 8  
O9VU79 PRELIMINARY; PRT; 1172 AA.

AC O9VU79;  
DT 01-MAY-2000 (TEMBLrel. 13, Created)  
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)  
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
DE CG10738 PROTEIN.  
GN CG10738.  
OS Drosophila melanogaster (Fruit fly).  
OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
OC Phrygane; Neoptera; Endopterygota; Diptera; Brachyera; Muscomorpha;  
OC Ephydroidea; Drosophilidae; Drosophila.  
OX NCBI\_Taxid=7227;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=BERKELEY;  
RX MEDLINE=20196006; Pubmed=10731132;  
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,  
RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,  
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,  
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,  
RA Brandon R.C., Rogers Y.-H.C., Blazer R.G., Champe M., Pfeiffer B.D.,  
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,  
RA Arroll J.F., Ashby A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,  
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,  
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,  
RA Borokva D., Botchan M.R., Bouck J., Brokstein P., Brotlier P.,  
RA Burks K.C., Busam D.A., Butler H., Cadiou E., Center A., Chandra I.,  
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,  
RA de Pablos P., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,  
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,  
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,  
RA Foster C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,  
RA Gload A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,  
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,  
RA Hostali D., Houston K.A., Howland T.J., Wei M.-H., Idegawa C.,  
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,  
RA Kimmel B.E., Kodire C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,  
RA Lascko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,  
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,  
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,  
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,  
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacle J.M.,  
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,  
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,  
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,  
RA Slier E., Spradling A.C., Stapleton M., Strong R., Sun E.,  
RA Sylvestre R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,  
RA Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissbach J.,  
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,  
RA Ye J., Yen K.-F., Zaveli J.S., Zhan W., Zhang G., Zhao Q., Zheng L.,  
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,  
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.,  
RT "The genome sequence of Drosophila melanogaster.";  
RL Science 287:2185-2195(2000).  
DR EMBL; A8003537; AAF49811.1; -.  
DR HSSP; Q02846; 1AWL.  
DR FLYBASE; FBgn0036368; CG10738.  
DR InterPro; IPR001828; ANF\_receptor.  
DR InterPro; IPR000719; Euk\_kinase.  
DR InterPro; IPR001054; Guanylt\_cyclase.  
DR Pfam; PF01094; ANF\_receptor; 2.  
DR Pfam; PF00211; guanylate\_cyc; 1.  
DR SMART; SM00044; CYCC; 1.  
DR PROSITE; PS00452; GUANYLATE\_CYCLASES\_1; 1.  
DR PROSITE; PS00125; GUANYLATE\_CYCLASES\_2; 1.

DR PROSITE; PS50011; PROTEIN\_KINASE\_DOM; 1.  
KW ATP-binding; Lyase; Transferase.  
SQ SEQUENCE 1172 AA; 13192 MW; C6450931F82A6D55 CRC64;

Query Match 79.58; Score 35; DB 5; Length 1172;  
Best Local Similarity 75.08; Pred. No. 1.4e+02;  
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPVGR 8  
I:|||||  
DB 434 GLRPVGR 441

RESULT 9  
O9C1M8 PRELIMINARY; PRT; 529 AA.

AC O9C1M8;  
DT 01-JUN-2001 (TEMBLrel. 17, Created)  
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)  
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)  
DE HEAT-INDUCED CATALASE.  
OS Pleurotus sajor-caju (Oyster mushroom).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
OC Agaricales; Pleurotaceae; Pleurotus.  
OX NCBI\_Taxid=50053;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Jeong M.-J., Park S.-C.;  
RT "Cloning of the catalase gene from Pleurotus sajor-caju.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF286097; AAK15159.1; -.  
DR HSSP; P15202; 1A4E.  
DR InterPro; IPR002226; Catalase.  
DR Pfam; PF00199; catalase; 1.  
DR PRINTS; PR00067; CATALASE.  
DR PRODOM; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE\_1; 1.  
DR PROSITE; PS00438; CATALASE\_2; UNKNOWN\_1.  
SQ SEQUENCE 529 AA; 59791 MW; A432BE6445B6AC13 CRC64;

Query Match 77.38; Score 34; DB 3; Length 529;  
Best Local Similarity 71.48; Pred. No. 97;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 2 IRPVGR 8  
I:|||||  
DB 298 LRPVGR 304

RESULT 10  
O9N6S1 PRELIMINARY; PRT; 139 AA.

AC O9N6S1;  
DT 01-OCT-2000 (TEMBLrel. 15, Created)  
DT 01-OCT-2000 (TEMBLrel. 15, Last sequence update)  
DT 01-OCT-2000 (TEMBLrel. 15, Last annotation update)  
DE L7535.11.  
GN L7535.11.  
OS Leishmania major.  
OC Eukaryota; Euklenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.  
OX NCBI\_Taxid=5664;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=RIEDELIN;  
RA Myler P.J.;  
RL Submitted (JAN-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AC005766; AAF28376.1; -.  
DR EMBL; AC005767; AAF27950.1; -.  
SQ SEQUENCE 139 AA; 15870 MW; 0E573F713CD9CBCE CRC64;

Query Match 75.0%; Score 33; DB 5; Length 139;  
 Best Local Similarity 75.0%; Pred. No. 37;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPGRF 8  
 |||||  
 16 GIRPGRF 23

RESULT 11  
 ID 033194 PRELIMINARY; PRT; 353 AA.  
 AC 033194.  
 DT 01-JAN-1998 (TREMBLrel. 05, Created)  
 DT 01-JAN-1998 (TREMBLrel. 05, Last sequence update)  
 DE 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
 GN HYPOTHETICAL. 37.0 KDA PROTEIN.  
 OS Mycobacterium tuberculosis.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Corynebacteriaceae; Mycobacteriaceae; Mycobacterium.  
 NCBI\_TaxID=1773;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-H37RV;  
 RA MEDLINE=98295987; PubMed=9634230;  
 RA Cole S.T., Brosch R., Parkhill J., Garnier T., Churcher C., Harris D.,  
 RA Gordon S.V., Eigmeier K., Gas S., Barry C.E. III, Tekala F.,  
 RA Badcock K., Basham D., Brown D., Chillingworth T., Connor R.,  
 RA Davies R., Devlin K., Feltwell T., Gentles S., Hamlin N., Holroyd S.,  
 RA Hornsby T., Jagels K., Krogh A., McLean J., Moule S., Murphy L.,  
 RA Oliver S., Osborne J., Quail M.A., Rajandream M.A., Rogers J.,  
 RA Rutter S., Seeger K., Skelton S., Squares S., Squares R.,  
 RA Sulston J.E., Taylor K., Whitehead S., Barrett B.G.;  
 RT "Deciphering the biology of Mycobacterium tuberculosis from the  
 RT complete genome sequence."  
 RL Nature 393:537-544(1998).  
 DR EMBL: Z98268; CAB10949.1; -.  
 DR Tuberculist; RV1692; -.  
 DR InterPro: IPR001454; HydroLase.  
 DR Pfam: PF00702; HydroLase; 1.  
 KW Hypothetical protein; Complete proteome.  
 SO SEQUENCE 353 AA; 37010 MW; 5DBE79E0040B5849 CRC64;

Query Match 75.0%; Score 33; DB 16; Length 353;  
 Best Local Similarity 75.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPGRF 8  
 |||||  
 107 GIRPGRF 114

RESULT 12  
 ID 066950 PRELIMINARY; PRT; 366 AA.  
 AC 066950.  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)  
 DT 01-AUG-1998 (TREMBLrel. 07, Last sequence update)  
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 GN TWITCHING MOTILITY PROTEIN PILT.  
 GN PILT OR AQ\_745.  
 OS Aquifex aeolicus.  
 OC Bacteria; Aquificales; Aquificaceae; Aquifex.  
 NCBI\_TaxID=63363;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-VF5;  
 RA MEDLINE=98196666; PubMed=9537320;  
 RA Deckert G., Warren P.V., Gaasterland T., Young W.G., Lenox A.L.,  
 RA Graham D.E., Overbeek R., Sneed M.A., Keller M., Anjay M., Huber R.,  
 RA Feldman R.A., Short J.M., Olson G.J., Swanson R.V.;

RT "The complete genome of the hyperthermophilic bacterium Aquifex  
 RT aeolicus."  
 RL Nature 392:353-358(1998).  
 DR EMBL: AE000705; AAC06903.1; -.  
 DR InterPro: IPR001687; ATP\_GTP\_A.  
 DR InterPro: IPR001482; GSPIL\_E.  
 DR Pfam: PF00437; GSPIL\_E; 1.  
 KW Complete proteome.  
 SO SEQUENCE 366 AA; 41205 MW; C043FBF7F80B41C1 CRC64;

Query Match 75.0%; Score 33; DB 16; Length 366;  
 Best Local Similarity 75.0%; Pred. No. 1e+02;  
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 GIRPGRF 8  
 |||||  
 87 GVRGGRF 94

RESULT 13  
 ID 050531 PRELIMINARY; PRT; 445 AA.  
 AC 050531.  
 DT 01-JUN-1998 (TREMBLrel. 06, Created)  
 DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
 DE 01-JUN-2001 (TREMBLrel. 17, Last annotation update)  
 GN PAD-DEPENDENT OXIDOREDUCTASE.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycetaceae; Streptomycetaceae; Streptomyces.  
 NCBI\_TaxID=1902;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (DEC-1997) to the EMBL/Genbank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parkhill J., Barrett B.G., Rajandream M.A.;  
 RL Submitted (NOV-1997) to the EMBL/Genbank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RX MEDLINE=97000351; PubMed=8843436;  
 RA Kienast H., Hopwood D.A.;  
 RA Kienast H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome."  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL: AL009204; CA15814.1; -.  
 DR InterPro: IPR001575; Oxid\_FAD\_bind.  
 DR Pfam: PF01565; FAD\_binding\_4; 1.  
 DR PROSITE: PS00862; OX2\_COVAL\_FAD; UNKNOWN\_1.  
 SO SEQUENCE 445 AA; 49067 MW; 8C0AEBF874E38CF5 CRC64;

Query Match 75.0%; Score 33; DB 2; Length 445;  
 Best Local Similarity 100.0%; Pred. No. 1.3e+02;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RPVGRF 8  
 |||||  
 248 RPVGRF 253

RESULT 14  
 ID 059714 PRELIMINARY; PRT; 511 AA.  
 AC 059714.  
 DT 01-AUG-1998 (TREMBLrel. 07, Created)

DT 01-AUG-1998 (Tremblrel. 07, last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)  
DE HYPOHETICAL 57.7 KDA PROTEIN C3B8.06 IN CHROMOSOME II.  
GN SPB3B8.06.  
OS Schizosaccharomyces pombe (Fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomycetes.  
OX NCBI\_Taxid=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=972;  
RA Lyne M., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.,  
RA Pohl T.;  
RA Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: IN C-TERMINAL TO YEAST YCR062W.  
DR EMBL; AL022244; CAA18295.1; -.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 33 53 POTENTIAL.  
FT TRANSMEM 59 79 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 332 352 POTENTIAL.  
FT TRANSMEM 379 399 POTENTIAL.  
FT DOMAIN 419 424 POLY-ALA.  
FT TRANSMEM 449 469 POTENTIAL.  
FT TRANSMEM 483 503 POTENTIAL.  
SQ SEQUENCE 511 AA; 57703 MW; BDDDE7650B2225E CRC64;

Query Match 75.0%; Score 33; DB 3; Length 511;  
Best Local Similarity 100.0%; Pred. No. 1.5e+02;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 3 RPVGRF 8  
|:|:|:|  
DB 95 RPVGRF 100

## RESULT 15

O9KSJ8 PRELIMINARY; PRT; 894 AA.  
AC O9KSJ8;  
DT 01-OCT-2000 (Tremblrel. 15, Created)  
DT 01-OCT-2000 (Tremblrel. 15, last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
DE DNA GYRASE, SUBUNIT A.  
GN VC1258.  
OS Vibrio cholerae.  
OC Bacteria; Proteobacteria; gamma subdivision; Vibrionaceae; Vibrio.  
OX NCBI\_Taxid=666;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=EL TOR N16961 / SEROTYPE O1;  
RX MEDLINE=20406833; Pubmed=10952301;  
RA Heidelberg J.F., Eisen J.A., Nelson W.C., Clayton R.A., Gwin M.L.,  
RA Dodson R.J., Haft D.H., Hickey E.K., Peterson J.D., Umayam L.A.,  
RA Gill S.R., Nelson K.E., Read T.D., Tettelin H., Richardson D.,  
RA Ermolaeva M.D., Vamathevan J., Bass S., Qin H., Dragoi I., Sellers P.,  
RA McDonald L., Utterback T., Fleischmann R.D., Nierman W.C., White O.,  
RA Salzberg S.L., Smith H.O., Colwell R.R., Mekalanos J.J., Venter J.C.,  
RA Fraser C.M.;  
RT "DNA sequence of both chromosomes of the cholera pathogen Vibrio  
RT cholerae.";  
RT Nature 406:477-483(2000).  
DR EMBL; AE004205; AAP94417.1; -.  
DR HSSP; P09097; IAB4.  
DR TIGR; VC1258; -.  
DR InterPro; IPR002205; DNA\_topoisomIV.  
DR Pfam; PF00521; DNA\_topoisomIV; 1.  
DR SMART; SM00434; TOP4c; 1.  
KW Complete proteome.  
SQ SEQUENCE 894 AA; 99025 MW; C7EA42056A08F57F CRC64;

Query Match 75.0%; Score 33; DB 16; Length 894;  
Best Local Similarity 71.4%; Pred. No. 2.7e+02;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
OY 1 GIRPVR 7  
|:|:|:|  
DB 734 GIRPVR 740

Search completed: September 13, 2002, 09:29:25  
Job time: 1070 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:36 ; Search time 399.68 Seconds  
(without alignments)  
2.223 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_14\_21

Perfect score: 38

Sequence: 1 IRPVGRFX 8

Scoring table:

BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

1: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1980.DAT:\*  
2: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1981.DAT:\*  
3: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1982.DAT:\*  
4: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1983.DAT:\*  
5: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1984.DAT:\*  
6: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1985.DAT:\*  
7: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1986.DAT:\*  
8: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1987.DAT:\*  
9: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1988.DAT:\*  
10: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1989.DAT:\*  
11: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1990.DAT:\*  
12: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1991.DAT:\*  
13: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1992.DAT:\*  
14: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1993.DAT:\*  
15: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1994.DAT:\*  
16: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1995.DAT:\*  
17: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1996.DAT:\*  
18: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1997.DAT:\*  
19: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1998.DAT:\*  
20: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA1999.DAT:\*  
21: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2000.DAT:\*  
22: /SIDSL/gcgdata/hold-geneseq/geneseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	97.4	8	22	AA846955
2	37	97.4	10	20	AAW95178
3	37	97.4	15	18	AAW31400
4	37	97.4	15	20	AAW97230
5	37	97.4	15	21	AAV48293
6	37	97.4	15	21	AAV49295
7	37	97.4	15	21	AAV49296
8	37	97.4	20	18	AAW31394
9	37	97.4	20	18	AAW31387
10	37	97.4	20	18	AAW31374
11	37	97.4	20	20	AAW97232

12	37	97.4	20	20	AAW97234
13	37	97.4	20	20	AAW97236
14	37	97.4	20	20	AAW95191
15	37	97.4	20	20	AAW95175
16	37	97.4	20	21	AAW95150
17	37	97.4	20	21	AAW10350
18	37	97.4	20	21	AAW10358
19	37	97.4	20	21	AAW10365
20	37	97.4	20	21	AAW10369
21	37	97.4	20	21	AAV49294
22	37	97.4	20	21	AAV49301
23	37	97.4	20	21	AAV49302
24	37	97.4	20	22	AAW62519
25	37	97.4	20	22	AAW62527
26	37	97.4	20	22	AAW62534
27	37	97.4	20	22	AAW62538
28	37	97.4	20	22	AAW90992
29	37	97.4	20	22	AAW90994
30	37	97.4	20	22	AAW90996
31	37	97.4	20	22	AAW90997
32	37	97.4	21	18	AAW31397
33	37	97.4	21	18	AAW31395
34	37	97.4	21	18	AAW31388
35	37	97.4	21	18	AAW31375
36	37	97.4	21	20	AAW97227
37	37	97.4	21	20	AAW87616
38	37	97.4	21	21	AAW95192
39	37	97.4	21	21	AAW10351
40	37	97.4	21	21	AAW10359
41	37	97.4	21	21	AAW10366
42	37	97.4	21	22	AAW62520
43	37	97.4	21	22	AAW62528
44	37	97.4	22	18	AAW31396
45	37	97.4	22	18	AAW31389

#### ALIGNMENTS

RESULT 1	
AA846955	standard; Protein; 8 AA.
AA846955;	
04-MAY-2001	(first entry)
Peptide PRP8 fragment.	
DE	
XX	
XX	
XX	
KW	GPR10; UHR-1; PRP receptor; prolactin-releasing peptide; pain;
KW	central nervous system disorder; autonomic regulation; analgesic;
XX	hypotensive; blood pressure.
OS	Unidentified.
XX	
XX	
FT	Key
FT	Modified-site
XX	Location/Qualifiers
XX	8
XX	/note- "C-terminal amide"
PN	W0200109182-A1.
PD	08-FEB-2001.
XX	
XX	03-AUG-2000; 2000MO-F100664.
XX	
XX	03-AUG-1999; 99US-0365756.
PR	20-MAR-2000; 2000US-0531567.
XX	
PA	(JUVVA-) JUVANTIA PHARMA LTD OY.
XX	
XX	Panula PAJ, Petrovaara A, Kalso E, Korpi E;
PI	
XX	
DR	WPI, 2001-182941/18.

XX C-terminal fragments of prolactin-releasing peptide useful for  
PT regulating autonomic functions and in the manufacture of a medicament  
PT for regulating blood pressure  
XX  
PS Claim 2; Page 10; 40pp; English.

XX This invention describes a novel C-terminal fragment (I) of an isolated  
CC prolactin-releasing peptide (PRP), referred to as PRP20 and having a  
CC sequence (S1). The invention also describes (1) a therapeutic composition  
CC (C1) comprising (I) or a C-terminal fragment of PRP referred to as PRP8  
CC and comprising a sequence (S2); (2) a diagnostic method based on antisera  
CC against PRP20 for identification of disorders involving the central  
CC nervous system, including those associated with pain or autonomic  
CC regulation, where specific antisera against the N-and/or C-terminal  
CC domains of PRP is used to identify alterations in PRP synthesis or  
CC levels; (3) a rat or human receptor encoded by a 1122 nucleotide sequence  
CC (S3), fully defined in the specification; (4) treating a person suffering  
CC from a disorder regulated by a receptor (II) encoded by a sequence of  
CC 1122 nucleotide sequence, fully defined in the specification, located in  
CC the central nervous system, by administering an agonist or antagonist to  
CC the receptor; and (5) treating blood pressure, by blocking of receptors  
CC of PRP or its C-terminal fragment GlylealrproValglyArgphe-NH<sub>2</sub> (S2).  
CC The products of the invention have analgesic and hypotensive activity.  
CC (1) is useful for regulating autonomic functions, such as increasing  
CC blood pressure. (1) is useful for treating pain, for manufacturing a  
CC medicament for regulating blood pressure, and for treating pain. Agonist  
CC and antagonist of (II) are useful for treating acute pain, inflammatory  
CC pain and neuropathic pain, for regulating autonomic functions and  
CC treating high blood pressure.

SQ Sequence 8 AA:

Query Match 97.4%; Score 37; DB 22; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6 4e+05;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPYGRF 7  
| | | | | | |  
Db 2 lrpvgrf 8

RESULT 2  
AAW95178  
ID AAW95178 standard; Protein; 10 AA.  
XX  
AC AAW95178;  
XX  
DT 10-MAR-1999 (first entry)  
XX  
DE Murine pituitary-derived ligand polypeptide antigenic epitope.  
XX  
KM Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; modulator; pituitary; central nervous system; pancreas;  
KM tIssue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; epitope.  
XX  
OS Mus sp.  
XX  
PN WO9849295-A1.  
XX  
PD 05-NOV-1998.  
XX  
PF 27-APR-1998; 98WO-JP01923.  
XX  
PR 28-APR-1997; 97JP-0109974.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fukusumi S, Hinuma S;  
XX  
DR WPI; 1999-009423/01.  
XX  
PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system; pituitary and  
PT pancreas, and for drug screening  
XX  
PS Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis, epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutant are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.

SQ Sequence 10 AA:

Query Match 97.4%; Score 37; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.21;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPYGRF 7  
| | | | | | |  
Db 4 lrpvgrf 10

RESULT 3  
AAW31400  
ID AAW31400 standard; Peptide; 15 AA.  
XX  
AC AAW31400;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Synthetic ligand 19P2-L31 peptide II.  
XX  
KM G-protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KM therapeutic agent; antigen.  
XX  
OS Synthetic.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
 PI Kawamata Y, Kitada C;  
 XX  
 DR WPI: 1997-363672/33.  
 XX  
 XX Lligand peptide for G protein-coupled receptor - acts by modulating  
 PT function in the central nervous system, pancreas and pituitary gland  
 XX  
 PS Example 43; Page 151; 258bp; English.  
 XX  
 CC This peptide contains the partial C-terminal sequence of the synthetic  
 CC 19p2 polypeptide 19p2-L31 which is capable of binding to a G  
 CC protein-coupled receptor protein. This peptide is used as an antigen to  
 CC prepare rabbit anti-bovine 19p2-L31 antibodies which are used in binding  
 CC assays. Pharmacological compositions containing this ligand may be used  
 CC as a pituitary function modulator, a central nervous system modulator  
 CC or a pancreatic function modulator. This ligand could have specific  
 CC applications as a prophylactic or therapeutic agent for dementia,  
 CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
 CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
 CC hyper- and polypagia, hyperlipidaemia, hypercholesterolaemia,  
 CC hyperglycaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
 CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
 CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
 CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
 CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
 CC and/or oligogalactia. Assays can also be developed to screen compounds  
 CC which are capable of altering the binding activity of the ligand  
 CC affecting activation of the G protein-coupled receptor protein.  
 CC  
 SQ Sequence 15 AA;  
 XX  
 XX  
 Query Match 97.4%; Score 37; DB 18; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 IRPVGRF 7  
 |||||  
 DB 9 IRPVGRF 15  
 XX  
 RESULT 4  
 AAM97230  
 ID AAM97230 standard; Peptide: 15 AA.  
 XX  
 AC AAM97230;  
 XX  
 DT 06-MAY-1999 (first entry)  
 XX  
 DE C-terminal ligand polypeptide derived antigen.  
 XX  
 KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cecogenesis;  
 KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;  
 KW pituitary adenomatosis; brain tumour; emmenorrhoea; autoimmune disease;  
 KW prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;  
 KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
 KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
 KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
 KW lrruprion mole; abortion; unfertility fetus; abnormal saccharometabolism;  
 KW abnormal lipidmetabolism; oxytocia; prolactin secretion.  
 XX  
 OS Synthetic.  
 XX  
 PN W09858962-A1.  
 PD 30-DEC-1998.  
 XX  
 PF 22-JUN-1998; 98WO-JP02765.  
 XX  
 PR 23-JUN-1997; 97JP-0165437.  
 XX  
 PA

XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX  
 PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
 XX  
 DR WPI: 1999-105614/09.  
 XX  
 XX Use of G protein-coupled receptor ligands - for modulating prolactin  
 PT secretion or placental function, e.g. for treating menopausal  
 PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
 XX  
 PS Example 43; Page 123; 241p; English.  
 XX  
 CC AAM97229-31 represent a ligand polypeptide derived fragments used to  
 CC produce antibodies. The specification describes an agent for modulating  
 CC prolactin secretion which comprises a ligand polypeptide or a salt, for  
 CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
 CC prolactin secretion can be used for treating or preventing  
 CC hypovarianism, gonocyst cecogenesis, menopausal syndrome, euthyroid or  
 CC hypometabolism. They can be used for promoting lactation in a domestic  
 CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
 CC secretion can be used for treating or preventing pituitary adenomatosis,  
 CC brain tumour, emmenorrhoea, autoimmune disease, prolactinoma,  
 CC infertility, impotence, amenorrhoea, galactorrhea, acromegaly,  
 CC Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright  
 CC syndrome, lymphoma, Sheehan syndrome or dyszoospermia. The inhibitory  
 CC agents can also be used as contraceptives. The agents for modulating  
 CC placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, irruption mole, abortion, unfertility  
 CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
 CC oxytocia.  
 CC  
 SQ Sequence 15 AA;  
 XX  
 XX  
 Query Match 97.4%; Score 37; DB 20; Length 15;  
 Best Local Similarity 100.0%; Pred. No. 0.32;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 OY 1 IRPVGRF 7  
 |||||  
 DB 9 IRPVGRF 15  
 XX  
 RESULT 5  
 AAY49293  
 ID AAY49293 standard; Peptide: 15 AA.  
 XX  
 AC AAY49293;  
 XX  
 DT 22-FEB-2000 (first entry)  
 XX  
 DE 19p2 ligand peptide fragment.  
 XX  
 KW Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
 KW pituitary; regulatory mechanism; central nervous system; pancreatitic.  
 XX  
 OS Homo sapiens.  
 XX  
 FT Key Location/Qualifiers  
 FT Modified-site 15 /note="C-terminal amide"  
 FT  
 XX  
 PN W09960112-A1.  
 PD 25-NOV-1999.  
 XX  
 PF 20-MAY-1999; 99WO-JP02650.  
 XX  
 PR 21-MAY-1998; 98JP-0140293.  
 XX  
 PA (TAKE ) TAKEDA CHEM IND LTD.  
 XX





Query Match 97.4% Score 37; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.32;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 IRPGRF 7  
| | | | |  
Db 9 lrpgrf 15

RESULT 8  
AAW31394  
ID AAW31394 standard; Peptide: 20 AA.  
XX  
AC AAW31394;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02431.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharides. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA:

Query Match 97.4% Score 37; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
Oy 1 IRPGRF 7  
| | | | |  
Db 14 lrpgrf 20

RESULT 9  
AAW31387  
ID AAW31387 standard; Peptide: 20 AA.  
XX  
AC AAW31387;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Rat type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Rat sp.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB: AAV02424.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 180; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the  
CC sequence represented in AAW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharides. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA:

Query Match 97.4% Score 37; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.43;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPVGRF 7  
| | | | | | |  
Db 14 lrpvgrf 20

RESULT 10  
AAW31374  
ID AAW31374 standard; Peptide: 20 AA.  
XX AAW31374;  
AC  
XX  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Bovine G protein-coupled receptor ligand peptide fragment 4.  
XX  
XX G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Bos taurus.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
XX 28-DEC-1995; 95UP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habeta Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kikada C;  
DR N-PSDB; AAW02397.  
XX  
XX  
PT ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
PS Claim 2; Page 161; 258pp; English.  
XX  
XX This sequence represents a peptide fragment of a novel bovine pituitary  
CC derived ligand corresponding to amino acid residues 34 to 53 of the  
CC sequence in AAW31368 and is used in an assay to monitor ligand binding  
CC to the G protein-coupled receptor protein. Pharmaceutical compositions  
CC containing this ligand may be used as a pituitary function modulator, a  
CC central nervous system modulator or a pancreatic function modulator.  
CC This ligand could have specific applications as a prophylactic or  
CC therapeutic agent for dementia, depression, hyperkinetic syndrome,  
CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
CC growth hormone secretory disease, hyper- and polyphagia,  
CC hypercholesterolaemia, hyperglyceridaemia, hyperlipidaemia,  
CC hyperprolactinaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, amyotrophic lateral sclerosis, acute  
CC myocardial infarction, spinocerebellar degeneration, bone fracture,  
CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility  
CC and/or oligosaccharia. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand thus  
CC affecting activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 97.4%; Score 37; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPVGRF 7  
| | | | | | |  
Db 14 lrpvgrf 20

RESULT 11  
AAW97232  
ID AAW97232 standard; peptide: 20 AA.  
XX  
XX AAW97232;  
AC  
XX  
XX  
DT 06-MAY-1999 (first entry)  
XX  
DE Bovine pituitary-derived ligand polypeptide fragment.  
XX  
XX  
XX Bovine pituitary-derived ligand; modulation; prolactin secretion;  
KM G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenests;  
KM menopausal syndrome; eutthyroid; hypometabolism; lactation;  
KM pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
KM prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KM acromegaly; Charli-Frommel syndrome; Argonz-del Castillo syndrome;  
KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KM contraceptive; placental function; choriocarcinoma; hydatid mole;  
KM interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;  
XX abnormal lipidmetabolism; oxytocia.  
XX  
OS Bos sp.  
XX  
PN WO9858962-A1.  
XX  
PD 30-DEC-1998.  
XX  
XX 22-JUN-1998; 98WO-JP02765.  
PF  
XX 23-JUN-1997; 97JP-0165437.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;  
PI  
DR WPI; 1999-105614/09.  
XX  
XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
XX  
PS Claim 3; Page 136; 241pp; English.  
XX  
XX The present sequence represents a bovine pituitary-derived ligand  
CC fragment. It is used in the course of the invention. The specification  
CC describes an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenests, menopausal  
CC syndrome, eutthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Charli-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, interruption mole, abortion, unfertilized fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.  
XX  
SQ Sequence 20 AA;

Query Match 97.4%; Score 37; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
|||||  
Db 14 lrpvgrf 20

## RESULT 12

AAW97234  
ID AAW97234 standard; peptide: 20 AA.

XX  
AC AAW97234:

DT 06-MAY-1999 (first entry)

XX  
DE Rat type ligand polypeptide fragment.

XX  
KW Rat type ligand: modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydaltid mole;  
KW Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.

XX  
OS Rattus sp.

PN MO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPI: 1999-105614/09.

PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy

XX  
PS Claim 3; Page 154; 241pp; English.

XX  
CC The present sequence represents a rat type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydaltid mole, Irruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX  
SQ Sequence 20 AA:

Query Match 97.4%; Score 37; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
|||||  
Db 14 lrpvgrf 20

## RESULT 13

AAW97236  
ID AAW97236 standard; peptide: 20 AA.

XX  
AC AAW97236:

DT 06-MAY-1999 (first entry)

XX  
DE Human type ligand polypeptide fragment.

XX  
KW Rat type ligand: modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmeniopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydaltid mole;  
KW Irruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.

XX  
OS Homo sapiens.

PN MO9858962-A1.

PD 30-DEC-1998.

PF 22-JUN-1998; 98WO-JP02765.

PR 23-JUN-1997; 97JP-0165437.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

DR WPI: 1999-105614/09.

PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy

XX  
PS Claim 3; Page 166; 241pp; English.

XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmeniopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydaltid mole, Irruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX  
SQ Sequence 20 AA:

Query Match 97.4%; Score 37; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
| | | | |  
Db 14 irpvgrf 20

## RESULT 14

AAW95191  
ID AAW95191 standard; peptide; 20 AA.

AC AAW95191;  
XX

DT 10-MAR-1999 (first entry)  
XX

DE Bovine pituitary-derived ligand polypeptide fragment.  
XX

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; bovine.  
XX

OS Bos sp.  
XX

PN WO9849295-A1.  
XX

PD 05-NOV-1998.  
XX

PF 27-APR-1998; 98WO-JP01923.  
XX

PR 28-APR-1997; 97JP-0109974.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fukusumi S, Hinuma S;  
XX

DR WPI; 1999-009423/01.  
XX

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

XX Example 19; Page 151; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of the  
CC polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. The present sequence represents a bovine genome-derived ligand  
CC polypeptide fragment which is similar to the murine ligand-polypeptide.  
XX

XX Sequence 20 AA;  
XX

Query Match 97.4%; Score 37; DB 20; Length 20;

Best Local Similarity 100.0%; Pred. No. 0.43;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 IRPVGRF 7  
| | | | |  
Db 14 irpvgrf 20

## RESULT 15

AAW95175  
ID AAW95175 standard; Protein; 20 AA.

AC AAW95175;  
XX

DT 10-MAR-1999 (first entry)  
XX

DE Murine pituitary-derived ligand polypeptide antigenic epitope.  
XX

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeldt-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; epitope.  
XX

OS Mus sp.  
XX

PN WO9849295-A1.  
XX

PD 05-NOV-1998.  
XX

PF 27-APR-1998; 98WO-JP01923.  
XX

PR 28-APR-1997; 97JP-0109974.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fukusumi S, Hinuma S;  
XX

DR WPI; 1999-009423/01.  
XX

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

XX Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeldt-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.  
XX

XX Sequence 20 AA;  
XX

Query Match 97.4%; Score 37; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 0.43;  
 Matches 7; Conservative 0; Mismatches 0; Indels 0;  
 Gaps 0;  
 QY 1 IRPVGRF 7  
 |||||  
 Db 14 IRPVGRF 20

Search completed: September 13, 2002, 09:18:36  
 Job time: 501 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:20:59 ; Search time 136.62 Seconds  
(without alignments)  
1,430 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_14\_21

Perfect score: 38

Sequence: 1 IRPVGRFX 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 10%

Listing first 45 summaries

Database : Issued\_Patents\_AA:\*

- 1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*
- 2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*
- 3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*
- 4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*
- 5: /cgn2\_6/ptodata/2/1aa/PCRTUS.COMB.pep:\*
- 6: /cgn2\_6/ptodata/2/1aa/Backfile1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	97.4	15	4	US-08-776-971-93
2	37	97.4	20	3	US-09-105-678A-34
3	37	97.4	20	3	US-09-105-678A-40
4	37	97.4	20	3	US-09-105-678A-46
5	37	97.4	20	4	US-08-776-971-8
6	37	97.4	20	4	US-08-776-971-50
7	37	97.4	20	4	US-08-776-971-64
8	37	97.4	20	4	US-08-776-971-98
9	37	97.4	20	4	US-09-421-208-34
10	37	97.4	20	4	US-09-421-208-40
11	37	97.4	20	4	US-09-421-208-46
12	37	97.4	21	3	US-09-105-678A-28
13	37	97.4	21	3	US-09-105-678A-35
14	37	97.4	21	3	US-09-105-678A-41
15	37	97.4	21	3	US-09-105-678A-47
16	37	97.4	21	4	US-08-776-971-9
17	37	97.4	21	4	US-08-776-971-51
18	37	97.4	21	4	US-08-776-971-65
19	37	97.4	21	4	US-09-421-208-28
20	37	97.4	21	4	US-09-421-208-35
21	37	97.4	21	4	US-09-421-208-41
22	37	97.4	21	4	US-09-421-208-47
23	37	97.4	22	3	US-09-105-678A-36
24	37	97.4	22	3	US-09-105-678A-42
25	37	97.4	22	3	US-09-105-678A-48
26	37	97.4	22	3	US-08-776-971-10
27	37	97.4	22	4	US-08-776-971-52

28	37	97.4	22	4	US-08-776-971-66	Sequence 66, Appl
29	37	97.4	22	4	US-08-776-971-73	Sequence 73, Appl
30	37	97.4	22	4	US-09-421-208-36	Sequence 36, Appl
31	37	97.4	22	4	US-09-421-208-42	Sequence 42, Appl
32	37	97.4	22	4	US-09-421-208-48	Sequence 48, Appl
33	37	97.4	31	3	US-09-105-678A-7	Sequence 7, Appl
34	37	97.4	31	3	US-09-105-678A-8	Sequence 8, Appl
35	37	97.4	31	3	US-09-105-678A-9	Sequence 9, Appl
36	37	97.4	31	3	US-09-105-678A-31	Sequence 31, Appl
37	37	97.4	31	3	US-09-105-678A-37	Sequence 37, Appl
38	37	97.4	31	3	US-09-105-678A-43	Sequence 43, Appl
39	37	97.4	31	4	US-09-172-353-4	Sequence 4, Appl
40	37	97.4	31	4	US-08-776-971-5	Sequence 5, Appl
41	37	97.4	31	4	US-08-776-971-47	Sequence 47, Appl
42	37	97.4	31	4	US-08-776-971-61	Sequence 61, Appl
43	37	97.4	31	4	US-08-776-971-97	Sequence 97, Appl
44	37	97.4	31	4	US-09-421-208-7	Sequence 7, Appl
45	37	97.4	31	4	US-09-421-208-8	Sequence 8, Appl

#### ALIGNMENTS

RESULT 1  
US-08-776-971-93  
; Sequence 93, Application US/08776971B  
; Patent No. 6228984  
; GENERAL INFORMATION:  
; APPLICANT: Hanuma, Shuji  
; Habata, Yugo  
; Kawamata, Yuji  
; Hosoya, Masaki  
; Fujii, Ryo  
; Fukusumi, Shoji  
; Kitada, Chieko  
; TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
; NUMBER OF SEQUENCES: 140  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FASTSEQ for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/776,971B  
; FILING DATE: 06-Feb-1997  
; CLASSIFICATION: <Unknown>  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/JP96/03821  
; FILING DATE: 28-DEC-1996  
; APPLICATION NUMBER: JP 7/343371  
; FILING DATE: 28-DEC-1995  
; APPLICATION NUMBER: JP 8/59419  
; FILING DATE: 15-MAR-1996  
; APPLICATION NUMBER: JP 8/211805  
; FILING DATE: 12-AUG-1996  
; APPLICATION NUMBER: JP 8/246573  
; FILING DATE: 18-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 47176  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 93:  
; SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93  
US-08-776-971-93

Query Match 97.4%: Score 37; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.085;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPYGRF 7  
|||||  
DB 9 IRPYGRF 15

RESULT 2  
US-09-105-678A-34  
Sequence 34, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 97.4%: Score 37; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPYGRF 7  
|||||  
DB 14 IRPYGRF 20

RESULT 3  
US-09-105-678A-40  
Sequence 40, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 97.4%: Score 37; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPYGRF 7  
|||||  
DB 14 IRPYGRF 20

RESULT 4  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA



ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 97.4%; Score 37; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
|||||||  
DB 14 IRPVGRF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
|||||||  
DB 14 IRPVGRF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
Db 14 IRPVGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
Db 14 IRPVGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 98:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
SEQUENCE DESCRIPTION: SEQ ID NO: 98:  
US-08-776-971-98

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
14 IRPVGRF 20

RESULT 9  
US-09-421-208-34  
Sequence 34, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421.208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-6440  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7

DB 14 IRPVGRF 20

RESULT 10  
US-09-421-208-40  
Sequence 40, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421.208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-6440  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
14 IRPVGRF 20

RESULT 11  
US-09-421-208-46  
Sequence 46, Application US/09421208  
Patent No. 6258561  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 97.4%; Score 37; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPYGRF 7  
Db 14 IRPYGRF 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product= "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product= "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product= "Gly-OH or Gly-Arg"  
US-09-105-678A-28

Query Match 97.4%; Score 37; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPYGRF 7  
Db 14 IRPYGRF 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 97.4%; Score 37; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPVGRF 7  
|||||||  
DB 14 IRPVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 97.4%; Score 37; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPVGRF 7  
|||||||

DB 14 IRPVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS: linear  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 97.4%; Score 37; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.12;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPVGRF 7  
|||||||  
DB 14 IRPVGRF 20

Search completed: September 13, 2002, 09:20:59  
Job time: 624 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:03 ; Search time 172.41 Seconds  
(without alignments)  
4.459 Million cell updates/sec

Title: us-09-446-543a-73\_COPY\_14\_21

Perfect score: 38

Sequence: 1 IRPVGRFX 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database :

1: PIR\_71:\*  
2: PIR1:\*  
3: PIR2:\*  
4: PIR3:\*  
5: PIR4:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	37	97.4	83	2 JC7607	prolactin-releasing
2	33	86.8	445	2 T35893	FAD-dependent oxid
3	33	86.8	511	2 T40334	hypothetical prote
4	32	84.2	149	2 AH2262	hypothetical prote
5	31	81.6	413	2 D12260	aminotransferase,
6	31	81.6	612	2 T36210	conserved hypothet
7	31	81.6	1551	2 AH3488	gramicidin s biosy
8	30	78.9	107	2 AC3538	hypothetical prote
9	30	78.9	154	2 T34825	hypothetical prote
10	30	78.9	173	2 S76779	hypothetical prote
11	30	78.9	198	2 A96520	hypothetical prote
12	30	78.9	274	2 D96840	hypothetical prote
13	30	78.9	292	2 T34529	hypothetical prote
14	30	78.9	292	2 T52257	CCR4-associated fa
15	30	78.9	313	2 B98339	integral membrane
16	30	78.9	373	2 T13024	probable protein k
17	30	78.9	374	2 S52645	probable 1-acyl-gl
18	30	78.9	376	2 S40471	DNA-directed RNA P
19	30	78.9	393	2 S25565	mitogen-activated
20	30	78.9	393	2 T47803	mitogen-activated
21	30	78.9	406	2 D84898	probable mitogen-a
22	30	78.9	435	2 H83307	probable porin PA2
23	30	78.9	457	2 A62728	polysaccharide bio
24	30	78.9	501	2 S66763	hypothetical prote
25	30	78.9	522	2 H97509	probable glycosyl
26	30	78.9	533	2 T20015	hypothetical prote
27	30	78.9	559	2 G83897	urocanase (urocana
28	30	78.9	601	2 H69274	probable long-chain
29	30	78.9	609	2 A99946	conserved hypothet

30	30	78.9	674	2 G70875	probable oxidoredu
31	30	78.9	751	2 F83080	hypothetical prote
32	29	76.3	170	2 S76067	hypothetical prote
33	29	76.3	172	2 T02239	protein BvJ15 - co
34	29	76.3	226	2 A65057	hypothetical prote
35	29	76.3	228	2 S64695	ribonuclease C (EC
36	29	76.3	239	2 A75060	hypothetical prote
37	29	76.3	289	2 A37209	thiosulfate sulfur
38	29	76.3	308	2 AC1053	probable membrane
39	29	76.3	315	2 T46156	hypothetical prote
40	29	76.3	362	2 G96735	probable proline-r
41	29	76.3	367	2 T04521	proline-rich prote
42	29	76.3	379	2 T45286	butyryl-CoA dehydr
43	29	76.3	397	2 G75182	DNA-directed RNA P
44	29	76.3	397	2 G71031	probable DNA-direc
45	29	76.3	472	2 T36529	probable two-compo

#### ALIGNMENTS

RESULT 1  
JC7607  
Prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Ishizuka, T.; Hosoya, A.;  
Biochem. Biophys. Res. Commun. 281, 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
A:Comment: This peptide induces arachidonic acid metabolite release from rat anterior  
release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PrRP  
A:Introns: 33/1

Query Match 97.4%; Score 37; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 0.72;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 IRPVGRF 7  
Db 46 IRPVGRF 52  
RESULT 2  
T35893  
FAD-dependent oxidoreductase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T35893  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajora, M.A.  
Submitted to the EMBL Data Library, November 1997  
A:Reference number: 221592  
A:Accession: T35893  
A:Status: preliminary; translated from GR/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-445 <OLI>  
A:Cross-references: EMBL:AL009204; PIDN:CA15814.1; GSPDB:GN00070; SCOPDB:SC9B10.24C  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOPDB:SC9B10.24C  
C:Superfamily: L-gulonolactone oxidase

Query Match 86.8%; Score 33; DB 2; Length 445;

Best Local Similarity 100.0%; Pred. No. 28;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 RPYGRF 7  
|||||  
Db 248 RPYGRF 253

RESULT 3  
T40334  
hypothetical protein SPBC3B8.06 - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40334  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.; Pohl, T.  
submitted to the EMBL Data Library, March 1998  
A:Reference number: Z21921  
A:Accession: T40334  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-511 <LYN>  
A:Cross-references: EMBL:AL022244; PIDN:CAA18295.1; GSPDB:GN00067; SPDB:SPBC3B8.06  
C:Genetics:  
A:Gene: SPDB:SPBC3B8.06  
A:Map position: 2

Query Match 86.8%; Score 33; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred. No. 32;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 2 RPYGRF 7  
|||||  
Db 95 RPYGRF 100

RESULT 4  
AH2262  
hypothetical protein alr3655 [imported] - *Anabaena* sp. (strain PCC 7120)  
C:Species: *Anabaena* sp.  
A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120  
C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH2262  
R:Nakano, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iritaguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.  
DNA Res. 8, 205-213, 2001  
A:Title: Complete genomic sequence of the filamentous nitrogen-fixing cyanobacterium *Anabaena*  
A:Reference number: AB1807; MIMD:2155285; PMID:11759840  
A:Accession: AH2262  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-149 <KUR>  
A:Cross-references: GB:BA000019; PIDN:BA075354.1; PID:q17132788; GSPDB:GN00179  
C:Genetics:  
A:Experimental source: strain PCC 7120  
A:Gene: alr3655

Query Match 84.2%; Score 32; DB 2; Length 149;  
Best Local Similarity 71.4%; Pred. No. 16;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;  
QY 1 IRPVGRF 7  
|||||  
Db 6 LRPVGRF 12

RESULT 5  
D72260  
aminotransferase, class V - *Thermotoga maritima* (strain MS8)  
C:Species: *Thermotoga maritima*  
C:Date: 11-Jun-1999 #sequence\_revision 11-Jun-1999 #text\_change 21-Jul-2000

C:Accession: D72260  
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwin, M.L.; Dodson, R.J.; Haft, D.H.; Hic  
Garrett, M.E.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson,  
C.M.  
Nature 399, 323-329, 1999  
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome  
A:Reference number: A72200; MIMD:99287316  
A:Accession: D72260  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-413 <ARN>  
A:Cross-references: GB:AE001791; GB:AE000512; MID:q4981929; PIDN:AAD36460.1; PID:q498  
A:Experimental source: strain MS8  
C:Genetics:  
A:Gene: TM1371  
C:Superfamily: nifs protein

Query Match 81.6%; Score 31; DB 2; Length 413;  
Best Local Similarity 85.7%; Pred. No. 71;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 IRPVGRF 7  
|||||  
Db 142 IRPVGRF 148

RESULT 6  
T36210  
conserved hypothetical protein SCE36.11c - *Streptomyces coelicolor* (fragment)  
C:Species: *Streptomyces coelicolor*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 20-Jun-2000  
C:Accession: T36210  
R:Oliver, K.; Harris, D.; Bentley, S.D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, May 1999  
A:Reference number: Z21601  
A:Accession: T36210  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-612 <OLI>  
A:Cross-references: EMBL:AL049763; PIDN:CAB42080.1; GSPDB:GN00070; SCOEDB:SCE36.11c  
C:Genetics:  
A:Gene: SCOEDB:SCE36.11c  
C:Superfamily: *Schizosaccharomyces pombe* hypothetical protein SPAC4H3.03c

Query Match 81.6%; Score 31; DB 2; Length 612;  
Best Local Similarity 85.7%; Pred. No. 1e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 IRPVGRF 7  
|||||  
Db 49 IRPVGRF 55

RESULT 7  
AH3488  
gramicidin S biosynthesis grs protein (EC 3.1.2.-) [imported] - *Brucella melitensis*  
C:Species: *Brucella melitensis*  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AH3488  
R:DelVecchio, V.G.; Kapatal, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanov,  
.; Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Let  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
A:Reference number: AD2522; PMID:11756688  
A:Accession: AH3488  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-1551 <KUR>  
A:Cross-references: GB:AE008917; PIDN:AA53075.1; PID:q17983937; GSPDB:GN00190  
A:Experimental source: strain 16M



C:Genetics:  
A:Gene: BMEI1894  
A:Map position: I  
C:Keywords: thiolester hydrolase

Query Match 81.6%; Score 31; DB 2; Length 1551;  
Best Local Similarity 85.7%; Pred. No. 2.7e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 IRPVGRF 7  
DB 1358 IRPVGSF 1364

RESULT 8  
AC3538  
hypothetical protein BMEI10229 [imported] - *Brucella melitensis* (strain 16M)  
C:Species: *Brucella melitensis*  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AC3538  
R:Delvecchio, V.G.; Kapetral, V.; Redkar, R.J.; Patra, G.; Mujer, C.; Los, T.; Ivanova,  
.: Mazur, M.; Goldsman, E.; Selkov, E.; Elzer, P.H.; Hagius, S.; O'Callaghan, D.; Letess  
Proc. Natl. Acad. Sci. U.S.A. 99,443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
A:Reference number: A03252; PMID:11756688  
A:Accession: AC3538  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-107 <KUR>  
A:Cross-references: GB:AE008918; PIDN:AA15470.1; PID:g17984372; GSPDB:GN00191  
A:Experimental source: strain 16M  
C:Genetics:  
A:Gene: BMEI10229  
A:Map position: II

Query Match 78.9%; Score 30; DB 2; Length 107;  
Best Local Similarity 83.3%; Pred. No. 30;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPYGRF 7  
DB 99 RPYGRF 104

RESULT 9  
T34825  
hypothetical protein SC2E9.14 SC2E9.14 - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34825  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, January 1998  
A:Reference number: Z21538  
A:Accession: T34825  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-154 <OL1>  
A:Cross-references: EMBL:AL021530; PIDN:CA16482.1; GSPDB:GN00070; SCOEDB:SC2E9.14  
A:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC2E9.14

Query Match 78.9%; Score 30; DB 2; Length 154;  
Best Local Similarity 83.3%; Pred. No. 43;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPYGRF 7  
DB 31 RPYGRF 36

RESULT 10  
S76779  
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
C:Species: *Synechocystis* sp.

A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76779  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O.; K. Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yanada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys  
s*.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76779  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-173 <KAN>  
A:Cross-references: EMBL:D90916; GB:AB001339; NID:g1653715; PIDN:BA18691.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 78.9%; Score 30; DB 2; Length 173;  
Best Local Similarity 83.3%; Pred. No. 49;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 RPYGRF 7  
DB 94 RPYGRF 99

RESULT 11  
A96520  
hypothetical protein T2J15.12 [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: A96520  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alon  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar,  
ansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim,  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Malt, R.; Marzia  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant *Arabidopsis*.  
A:Reference number: A86141; MUID:21016719  
A:Accession: A96520  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-198 <STO>  
A:Cross-references: GB:AE005173; NID:g10645475; PIDN:AA621589.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: T2J15.12  
A:Map position: 1

Query Match 78.9%; Score 30; DB 2; Length 198;  
Best Local Similarity 83.3%; Pred. No. 56;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 IRPVGR 6  
DB 127 VRPVGR 132

RESULT 12  
D96840  
hypothetical protein F23A5.13 [imported] - *Arabidopsis thaliana*  
C:Species: *Arabidopsis thaliana* (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D96840

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; Hansen, N.F.; Hughes, B.; Huizart, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maltli, R.; Marzall, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, K.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719  
A:Accession: D96840  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-274 <STO>  
A:Cross-references: GB:AE05173; NID:96503290; PIDN:AAF14666.1; GSPDB:GN00141  
A:Gene: F23A5.13  
A:Map position: 1

Query Match 78.9% Score 30; DB 2; Length 274;  
Best Local Similarity 71.4% Pred. No. 77;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
:||:|  
Db 48 VRPVGTF 54

RESULT 13  
T34529  
hypothetical protein DKFZp434M151.1 - human  
C:Species: Homo sapiens (man)  
C:Date: 29-Oct-1999 #sequence\_revision 29-Oct-1999 #text\_change 29-Oct-1999  
C:Accession: T34529  
A:Reference number: 221539  
R:Bioecker, H.; Beecher, M.; Brandt, P.; Mewes, H.W.; Gassenhuber, J.; Wiemann, S.  
Submitted to the Protein Sequence Database, October 1999  
A:Accession: T34529  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-292 <BLD>  
A:Cross-references: EMBL:AL122045  
A:Experimental source: adult testis; clone DKFZp434M151  
C:Genetics:  
A:Note: DKFZp434M151.1

Query Match 78.9% Score 30; DB 2; Length 292;  
Best Local Similarity 57.1% Pred. No. 82;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
:||:|  
Db 48 VRPVGTF 54

RESULT 14  
T52257  
CCRA-associated factor 1 [Imported] - human  
C:Species: Homo sapiens (man)  
C:Date: 20-Oct-2000 #sequence\_revision 20-Oct-2000 #text\_change 20-Oct-2000  
C:Accession: T52257  
R:Fidler, C.; Waincoat, J.S.; Boulwood, J.  
Genomics 56, 134-136, 1999  
A:Title: The human POP2 gene: identification, sequencing, and mapping to the critical re  
A:Reference number: 226006; MUID:9913686  
A:Accession: T52257  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-292 <FTD>  
A:Cross-references: EMBL:AF053318; PIDN:AAD02685.1

C:Genetics:  
A:Gene: POP2  
A:Map position: 5

Query Match 78.9% Score 30; DB 2; Length 292;  
Best Local Similarity 57.1% Pred. No. 82;  
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
:||:|  
Db 48 VRPVGTF 54

RESULT 15  
B98339  
Integral membrane protein (D85027) [Imported] - Agrobacterium tumefaciens (strain C58  
C:Species: Agrobacterium tumefaciens  
C:Date: 22-Oct-2001 #sequence\_revision 22-Oct-2001 #text\_change 11-Jan-2002  
C:Accession: B98339  
R:Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldm  
Science 294, 2323-2328, 2001  
A:Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium  
A:Reference number: A97359; PMID:11743194  
A:Accession: B98339  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-313 <KUR>  
A:Cross-references: GB:AE007870; PIDN:AAK90236.1; PID:915160251; GSPDB:GN00170  
C:Genetics:  
A:Gene: AGR\_L3320  
A:Map position: linear chromosome  
C:Superfamily: inner membrane protein upPA

Query Match 78.9% Score 30; DB 2; Length 313;  
Best Local Similarity 83.3% Pred. No. 88;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGR 6  
||:|  
Db 17 IRPIGR 22

Search completed: September 13, 2002, 09:24:04  
Job time: 779 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:49 ; Search time 80.21 Seconds  
(without alignments)  
3.862 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_14\_21  
Perfect score: 38  
Sequence: 1 IRPVGRFX 8

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues  
Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Length	ID	Description
1	37	97.4	83	PRRP_RAT
2	37	97.4	87	PRRP_HUMAN
3	37	97.4	98	PRRP_BOVIN
4	31	81.6	250	LINC_PSEPA
5	31	81.6	413	CSD_THEMA
6	30	78.9	492	CNT8_HUMAN
7	30	78.9	376	MPK5_ARATH
8	30	78.9	393	PPA2_THREE
9	30	78.9	480	GLG1_RHIME
10	29	76.3	130	RL3_PIG
11	29	76.3	224	YGC1_ECOLI
12	29	76.3	289	THTR_CHICK
13	29	76.3	330	AR72_HUMAN
14	29	76.3	331	AR73_HUMAN
15	29	76.3	397	PPA2_HUMAN
16	29	76.3	397	PPA2_PIRAB
17	29	76.3	491	PPA2_PIRAB
18	29	76.3	492	CAT2_ORYSA
19	29	76.3	496	CAT3_CUCPE
20	29	76.3	619	HEX3_YEAST
21	29	76.3	1693	POLN_HEYBU
22	28	73.7	153	MAUF_METHY
23	28	73.7	283	PSTA_RHITO
24	28	73.7	285	MAUF_METHY
25	28	73.7	292	YAAD_RHISN
26	28	73.7	299	MGRB_LACIA
27	28	73.7	300	YR33_THREE
28	28	73.7	313	YHCL_BACSU
29	28	73.7	376	MPK4_ARATH
30	28	73.7	378	METK_BUCAI
31	28	73.7	380	METK_RICPR
32	28	73.7	382	METK_RICPR
33	28	73.7	383	METK_ECOLI

34	28	73.7	384	METK_HAETN	P43762 haemophilus
35	28	73.7	384	METK_PASMU	P57897 pasteurella
36	28	73.7	396	METK_TREPA	O83772 treponema p
37	28	73.7	397	METK_STAU	P50307 staphylococ
38	28	73.7	399	METK_BACHD	O9c7q9 bacillus ha
39	28	73.7	399	METK_LACIA	O9c6e0 lactococcus
40	28	73.7	400	METK_BACSU	P54419 bacillus su
41	28	73.7	433	ZUO1_YEAST	P33527 saccharomyc
42	28	73.7	512	ATPA_RHOBL	P05439 rhodospirilo
43	28	73.7	526	CATP_BRARE	O9c7q2 brachydanio
44	28	73.7	847	MDOH_ECOLI	P33137 escherichia
45	28	73.7	1057	ANPA_MOUSE	P18293 mus musculu

## ALIGNMENTS

RESULT 1

ID	PRRP_RAT	STANDARD:	PRT:	83 AA.
AC	P81278;			
DT	30-MAY-2000 (Rel. 39, Created)			
DT	30-MAY-2000 (Rel. 39, Last sequence update)			
DT	01-MAR-2002 (Rel. 41, Last annotation update)			
DE	Proactin-releasing peptide precursor (PRP) (proactin-releasing			
DE	hormone) [Contains: Proactin-releasing peptide PRRP1; Proactin-			
DE	releasing peptide PRRP20].			
GN	PRH.			
OS	Rattus norvegicus (Rat).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Rattus.			
OX	NCBI_Taxid=10116;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	TISSUE=Brain;			
RX	MEDLINE=96266781; Pubmed=9607765;			
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,			
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,			
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.;			
RA	"A proactin-releasing peptide in the brain.";			
RT	Nature 393:272-276(1998).			
RL	[2]			
RN	TISSUE SPECIFICITY.			
RP	Pubmed=10498338;			
RX	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,			
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,			
RA	Funino Y., Fujino M.;			
RT	"Tissue distribution of proactin-releasing peptide (PRP) and its			
RT	receptor.";			
RL	Regul. Pept. 83:1-10(1999).			
CC	-I- FUNCTION: Stimulates prolactin (PRL) release and regulates the			
CC	expression of prolactin through its receptor GPR10. May stimulate			
CC	lactotrophs directly to secrete PRL.			
CC	-I- TISSUE SPECIFICITY: Widely expressed, with highest levels in			
CC	medulla oblongata and hypothalamus.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration			
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -			
CC	the European Bioinformatics Institute. There are no restrictions on its			
CC	use by non-profit institutions as long as its content is in no way			
CC	modified and this statement is not removed. Usage by and for commercial			
CC	entities requires a license agreement (see <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/announce/</a>			
CC	or send an email to <a href="mailto:license@sib-sib.ch">license@sib-sib.ch</a> ).			
CC	-----			
DR	EMBL; AB015418; BAA29026.1;			
KW	Hormone; Amidation; Signal; Cleavage on pair of basic residues.			
FT	SIGNAL	1	21	BY SIMILARITY.
FT	PEPTIDE	22	52	PROLACTIN-RELEASING PEPTIDE PRRP31.
FT	PEPTIDE	33	52	PROLACTIN-RELEASING PEPTIDE PRRP20.
FT	MOD. RES.	52	52	AMIDATION (G-53 PROVIDE AMIDE GROUP).
SO	SEQUENCE	83 AA;	9215 MW;	DOC75A264EBEAF29 CRC64;



RC STRAIN-UT26;  
RA MEDLINE-94252977; PubMed-7515041;  
RA Nagata Y., Ohtomo R., Miyachi K., Fukuda M., Yano K., Takagi M.;  
RT "Cloning and sequencing of a 2,5-dichloro-2,5-cyclohexadiene-1,4-diol  
dehydrogenase gene involved in the degradation of gamma-  
RT hexachlorocyclohexane in *Pseudomonas paucimobilis*.";  
RL Bacteriol. 176:3117-3125(1994).  
CC -1- FUNCTION: DEGRADATION OF 2,5-DICHLORO-2,5-CYCLOHEXADIENE-1,4-DIOL  
CC (2,5-DIOL) INTO 2,5-DICHLOROHYDROQUINONE (2,5-DIOL).  
CC -1- PATHWAY: DEGRADATION OF GAMMA-HEXACHLOROCYCLOHEXANE.  
CC -1- SIMILARITY: BELONGS TO THE SHORT-CHAIN DEHYDROGENASES/REDUCTASES  
CC (SDR) FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: D14595; BAA03444.1; -  
DR HSSP: P19992; 1HC.  
DR InterPro: IPR002198; ADH\_short.  
DR Pfam: PF00106; adh\_short; 1.  
DR PRINTS: PR00080; SDRFAMILY.  
DR PROSITE: PS00061; ADH\_SHORT; 1.  
KM Aromatic hydrocarbons catabolism: Oxidoreductase; NAD.  
FT NP\_BIND 9 34 NAD (BY SIMILARITY).  
FT ACT\_SITE 154 154 BY SIMILARITY.  
SQ SEQUENCE 250 AA; 25644 MW; PFCICAEB47DF89D CRC64;  
  
Query Match 81.6%; Score 31; DB 1; Length 250;  
Best Local Similarity 71.4%; Pred. No. 13;  
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 IRPVGRF 7  
DB 209 IHPICRF 215  
  
RESULT 5  
CSD\_THEME STANDARD; PRT; 413 AA.  
ID O9X191;  
AC O9X191;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Probable cysteine desulfurase (EC 4.4.1.-).  
GN CSD OR TM1371.  
OS Thermotoga maritima.  
OC Bacteria; Thermotogales; Thermotoga.  
OX NCBI\_TaxID=2336;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-MSB8 / DSM 3109;  
RX MEDLINE-99287316; PubMed-10360571;  
RA Nelson K.E., Clayton R.A., Gill S.R., Gwin M.L., Dodson R.J.,  
RA Haft D.H., Hickey E.K., Peterson J.D., Nelson W.C., Ketchum K.A.,  
RA McDonald L., Utterback T.R., Malek J.A., Linher K.D., Garrett M.M.,  
RA Stewart A.M., Cotton M.D., Pratt M.S., Phillips C.A., Richardson D.,  
RA Heidelberg J., Sutton G.G., Fleischmann R.D., Eisen J.A., White O.,  
RA Salzberg S.L., Smith H.O., Venter J.C., Fraser C.M.;  
RT "Evidence for lateral gene transfer between Archaea and Bacteria from  
RL genome sequence of *Thermotoga maritima*.";  
RL Nature 399:323-329(1995).  
CC -1- FUNCTION: CATALYZES THE REMOVAL OF ELEMENTAL SULFUR AND SELENIUM  
CC ATOMS FROM L-CYSTEINE, L-CYSTEINE, L-SELENOCYSTEINE, AND L-  
CC SELENOCYSTEINE TO PRODUCE L-ALANINE (BY SIMILARITY).  
CC -1- COFACTOR: PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO CLASS-V OF PYRIDOXAL-PHOSPHATE-DEPENDENT  
CC AMINOTRANSFERASES. CSD SUBFAMILY.

CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AE001791; AAD36460.1; -  
DR HSSP: P77444; 1CON.  
DR TIGR: TM1371; -  
DR InterPro: IPR00192; AminoTransf\_class.V.  
DR Pfam: PF00266; aminotran\_5; 1.  
DR PROSITE: PS00595; AA\_TRANSF\_CLASS\_5; FALSE\_NEG.  
KM Lysase: Pyridoxal phosphate; Complete proteome.  
FT NP\_BIND 223 223 PYRIDOXAL PHOSPHATE (BY SIMILARITY).  
SQ SEQUENCE 413 AA; 46748 MW; 09286633F984F0 CRC64;  
  
Query Match 81.6%; Score 31; DB 1; Length 413;  
Best Local Similarity 85.7%; Pred. No. 22;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
  
OY 1 IRPVGRF 7  
DB 142 IRPSCRF 148  
  
RESULT 6  
CSD\_THEME STANDARD; PRT; 292 AA.  
ID CNT8\_HUMAN  
AC G9UFE9; G95709;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE CCR4-NOT transcription complex, subunit 8 (CAF1-like protein)  
DE (CALIFP).  
GN CNOT8 OR CALIF OR POP2.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-99156866; PubMed-10036195;  
RA Fidler C., Waincoat J.S., Boulwood J.;  
RT "The human POP2 gene: identification, sequencing, and mapping to the  
RL critical region of the 5q- syndrome.";  
RL Genomics 56:134-136(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-20105461; PubMed-10637334;  
RA Albert T.K., Lemaire M., van Herkum N.L., Gentz R., Collart M.A.,  
RA Timmers H.T.M.;  
RT "Isolation and characterization of human orthologs of yeast CCR4-NOT  
RL complex subunits.";  
RL Nucleic Acids Res. 28:809-817(2000).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Testis;  
RX MEDLINE-21154917; PubMed-11230166;  
RA Wiemann S., Well B., Wellenreuther R., Gassenhuber J., Glasl S.,  
RA Anderson W., Boecker M., Bloecher H., Bauersachs S., Blum H.,  
RA Lauber J., Duesterhoeft A., Beyer A., Koehler K., Strack N.,  
RA Mewes H.-W., Ottenwelder B., Obermayer B., Tampe J., Heubner D.,  
RA Wambutt R., Korn B., Klein M., Poustka A.;  
RT "Towards a catalog of human genes and proteins: sequencing and  
RL analysis of 500 novel complete protein coding human cDNAs.";  
RL Genome Res. 11:422-435(2001).  
CC -1- FUNCTION: UNKOWN TRANSCRIPTION FACTOR REQUIRED FOR A DIVERSE  
CC SET OF PROCESSES. IT IS A COMPONENT OF THE CCR4 COMPLEX INVOLVED  
CC IN THE CONTROL OF GENE EXPRESSION.

```

CC -1 SUBCELLULAR LOCATION: Nuclear.
CC -1 SIMILARITY: BELONGS TO THE CAP1 FAMILY.
-----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation.
CC the European Bioinformatics Institute. There are no restrictions on
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AF053318; AAD02685.1; -
DR EMBL; AF180476; AAF29680.1; -
DR EMBL; AL122045; CAB59181.1; -
DR MIM: 603731; -
KW Transcription regulation; Repressor; Nuclear protein.
FT CONFLICT 32 32 L -> P (IN REF. 1).
FT CONFLICT 182 182 F -> L (IN REF. 1).
FT CONFLICT 189 189 F -> S (IN REF. 1).
FO SEQUENCE 292 AA; 33540 MW; 81027A966E51A6B CRC64;

```

Query Match	78.98;	Score 30;	DB 1;	Length 292;
Best Local Similarity	57.18;	Pred. NO. 26;		
Matches	4;	Conservative	2;	Mismatches 1;
				Indels 0;
				Gaps 0;

Qy	1	IRPVGRF	7
	:	:	
Db	48	VRPIGEF	54

RESULT	7	
MPK5_ATH		
ID	MPK5_ATH	STANDARD;
		PRT;
		376 AA.

DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 30-MAR-2000 (Rel. 39, Last annotation update)  
DE Mitogen-activated protein kinase homolog 5 (EC 2.7.1.-) (MAP kinase 5) (AKMPK5).

OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota: Viridiplantae: Streptophyta: Embryophyta: Tracheophyta:  
OC Spermatophyta: Magnoliophyta: eudicotyledons, core eudicots: Rosidae:  
OC eurosids II: Brassicales: Brassicaceae: Arabidopsis.  
NCBI TaxID=3702;

RP SEQUENCE FROM N.A.  
RC SPRAIN-CV, COLUMBIA,  
RX MEDLINE-94109583; PubMed-8282107;  
RA Mizoguchi T., Hayashida N., Yamaguchi-Shinozaki K., Kamada H.,  
RA Shinozaki K.;

RL FEBS Lett. 336:440-444(1993)  
CC -1- ENZYME REGULATION: ACTIVATED BY TYROSINE AND THREONINE  
CC PHOSPHORYLATION (BY SIMILARITY).  
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES  
CC MAP KINASE SUBFAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

DR EMBL: D21841; BAA004868.1; -.  
DR HSSP: P27703; 1ERK.  
DR InterPro: IPR000719; Euk\_pkinase.  
DR InterPro: IPR003527; MAP\_kin.  
DR InterPro: IPR002290; Ser\_thr\_pkinase  
DR Pfam: PF00069; pkinase; 1.

DR	SMART:SM00220; S. TARK; 1.				
DR	PROSITE; PS00351; MKK; 1.				
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.				
DR	PROSITE; PSS0011; PROTEIN_KINASE_DOM; 1.				
DR	PROSITE; PS00108; PROTEIN_KINASE_SF; 1.				
KW	transferase; Serine/threonine-protein kinase; ATP-binding;				
KW	phosphorylation; Multigene family.				
KM	DOMAIN	43	329		PROTEIN KINASE.
FT	NP_BIND	49	57		ATP (BY SIMILARITY).
FT	BINDING	72	72		ATP (BY SIMILARITY).
FT	ACT_SITE	169	169		BY SIMILARITY.
FT	MOD_RES	201	201		PHOSPHORYLATION (ACTIVATES THE
FT					(BY SIMILARITY).
FT	MOD_RES	203	203		PHOSPHORYLATION (ACTIVATES THE
FT					(BY SIMILARITY).
FT					(BY SIMILARITY).
SO	SEQUENCE	376 AA;	43120 MM;	0631F9E9B654A22EB	CRC64;

Query Match	78.9%	Score 30	DB 1	Length 376
Best Local Similarity	83.3%	Pred. No. 33		
Matches	5	Conservative	1	Mismatches 0; Indels 0; Gaps 0;

Qy	1	IRPVGR	6
	111	:	11
Db	46	IRPIGR	51

RESULT	8
RPA2_THECE	
ID	RPA2_THECE
STANDARD;	
PRT;	393 AA

DT 01-DEC-1992 (Rel. 24, Created)  
DT 01-JUL-1993 (Rel. 26, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE DNA-directed RNA polymerase subunit A" (EC 2.7.7.6)  
DE proc3

OS Thermococcus celer;  
OC Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Thermococcus  
NCBI\_TaxID=2264;  
[1]  
RN  
RP SEQUENCE FROM N.A.

RA MDLINE=93027167; Pubmed=1408768;  
 RX Klank H.-P., Schwass V., Lottspeich F., Zillig W.;  
 RT "Nucleotide sequence of the genes encoding the three largest subunits  
 of the DNA-dependent RNA polymerase from the archaeum *Thermococcus*  
 celer.";

RN [2]  
RP SEQUENCE OF 195-393 FROM N.A.  
RC STRAIN=DSM 2476;  
RX MEDLINE=9205139; PubMed=1840672;  
RA Klek H.-P., Schwaas V., Zillig W.;

RT equivalent ribosomal proteins from the archaeum *Thermococcus celer*.",  
 RL Nucleic Acids Res. 19:6047-6047(1991).  
 CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION  
 CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS  
 CC SUBSTRATES.

CC - [RNA](N).  
CC -  
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collab  
CC between the Swiss Institute of Bioinformatics and the EMBL outst

CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR	EMBL; X67313; CAA47724.1; -.
DR	EMBL; X60305; CAA42846.1; -.

DR PIR: S25565; S25565.  
 DR InterPro: IPR002879; RNA\_pol\_A2.  
 AC Pfam: PF01854; RNA\_pol\_A2; 1.  
 KW Transferase; Transcription; DNA-directed RNA polymerase.  
 SO SEQUENCE 393 AA; 43888 MW; 23BFAD738E10792E CRC64;

Query Match 78.9%; Score 30; DB 1; Length 393;  
 Best Local Similarity 83.3%; Pred. No. 34;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGR 6  
 DB 320 INPLGR 325

RESULT 9  
 ID GLGI\_RHIME STANDARD; PRT; 480 AA.  
 AC P58393;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Glycogen synthase 1 (EC 2.4.1.21) (Starch [bacterial glycogen] synthase 1).  
 GN GLGA1 OR R02846 OR SMC03924.  
 OS Rhizobium meliloti (Sinorhizobium meliloti).  
 CC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;  
 CC Rhizobiaceae; Sinorhizobium.  
 OX NCBI\_TaxID=382;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=21396507; PubMed=11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J., Bolstad P., Becker A., Boutry M., Cadieu E., Dreano S., Gloux S., Goerle T., Goffeau A., Kahn D., Kiss E., Lelaure V., Masny D., Pohl T., Portetalle D., Puehler A., Purnelle B., Ramsperger U., Renard C., Thebault P., Vandenbol M., Weidner S., Galibert F.;  
 RT "Analysis of the chromosome sequence of the legume symbiont Sinorhizobium meliloti strain 1021."  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.  
 CC -1- CATALYTIC ACTIVITY: ADP-glucose + ((1,4)-alpha-D-glucosyl)(N) -> ADP + ((1,4)-alpha-D-glucosyl)(N+1).  
 CC -1- PATHWAY: Glycogen biosynthesis; second step.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE FAMILY.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: AL591792; CAC47425.1; -  
 KW Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 KW Complete proteome.  
 FT BINDING 15 ADP-GLUCOSE (BY SIMILARITY).  
 SO SEQUENCE 480 AA; 51408 MW; 84C584F6E0564097 CRC64;

Query Match 78.9%; Score 30; DB 1; Length 480;  
 Best Local Similarity 83.3%; Pred. No. 42;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPYGRF 7  
 DB 55 KPYGRF 60

RESULT 10

ID R3\_PIG STANDARD; PRT; 130 AA.  
 AC Q29293;  
 DT 15-DEC-1998 (Rel. 37, Created)  
 DT 15-DEC-1998 (Rel. 37, Last sequence update)  
 DT 15-DEC-1998 (Rel. 37, Last annotation update)  
 DE 60S ribosomal protein L3 (Fragment).  
 GN RPL3.  
 OS Sus scrofa (pig).  
 CC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Small intestine;  
 RX MEDLINE=96327607; PubMed=8672129;  
 RT Wintrobe A.K., Fredholm M., Davies W.;  
 RT "Evaluation and characterization of a porcine small intestine cDNA library: analysis of 839 clones."  
 RL Mann. Genome 7:509-517(1996).  
 CC -1- FUNCTION: THE L3 PROTEIN IS A COMPONENT OF THE LARGE SUBUNIT OF CYTOPLASMIC RIBOSOMES.  
 CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
 CC -1- SIMILARITY: BELONGS TO THE L3P FAMILY OF RIBOSOMAL PROTEINS.

-----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL: F14648; CA23176.1; -  
 DR InterPro: IPR000597; Ribosomal\_L3.  
 DR Pfam: PF00297; Ribosomal\_L3; 1.  
 DR ProDom: PD001374; Ribosomal\_L3; 1.  
 DR PROSITE: PS00474; RIBOSOMAL\_L3; 1.  
 KW Ribosomal protein.  
 FT NON\_TER 1 1  
 FT NON\_TER 130 130  
 SO SEQUENCE 130 AA; 14766 MW; ACA48F9D3437A4BF CRC64;

Query Match 76.3%; Score 29; DB 1; Length 130;  
 Best Local Similarity 71.4%; Pred. No. 19;  
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGR 7  
 DB 111 INPLGRF 117

RESULT 11  
 ID YGCI\_ECOLI STANDARD; PRT; 224 AA.  
 AC Q46898;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein ygcI precursor.  
 GN YGCI OR B2757.  
 OS Escherichia coli.  
 CC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 CC Escherichia.

OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=92728503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,

RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of *Escherichia coli* K-12.";  
 RL Science 277:1453-1474(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: U29579; AAA69267.1; ALT\_INIT.  
 DR EMBL: AB000359; AAC75799.1; ALT\_INIT.  
 DR EcoGene: EGJ3116; YgcI.  
 KM Hypothetical protein; Signal; Complete proteome.  
 FT SIGNAL 1 17  
 FT CHAIN 18 224  
 FT SEQUENCE 224 AA; 25209 MW; 1C42CC009B317D68 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 224;

Best Local Similarity 83.3%; Pred. No. 33;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RPVGRF 7  
 || ||  
 Db 25 RPTGRF 30

RESULT 12  
 THIR-CHICK STANDARD; PRT; 289 AA.  
 ID THIR-CHICK  
 AC P25324;  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Thiosulfate sulfurtransferase (EC 2.8.1.1) (Rhodanese).  
 GN TST.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauromorpha; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OC NCBI\_TaxID=9031;  
 RN [1]  
 RP TISSUE=Liver;  
 RC MEDLINE=9113289; PubMed=2275748;  
 RA Kohanski R.A., Heinrichson R.L.;  
 RT "Primary structure of avian hepatic rhodanese.";  
 RL J. Protein Chem. 9:369-377(1990).  
 CC -1- FUNCTION: FORMATION OF IRON-SULFUR COMPLEXES AND CYANIDE  
 CC DETOXIFICATION.  
 CC -1- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- TISSUE SPECIFICITY: FOUND IN NUMEROUS TISSUES.  
 CC -1- DOMAIN: THE STRUCTURE CONSISTS OF 2 DOMAINS OF VERY SIMILAR  
 CC CONFORMATION, SUGGESTING A COMMON EVOLUTIONARY ORIGIN. HOWEVER,  
 CC THE SEQUENCES OF THE 2 DOMAINS ARE VERY DIFFERENT.  
 CC -1- SIMILARITY: BELONGS TO THE RHODANESE FAMILY.  
 DR PIR: A37209; A37209.  
 DR HSSP: P00586; IRHS.  
 DR InterPro: IPR001307; Rhodanese.  
 DR InterPro: IPR001763; Rhodanese\_domain.  
 DR Pfam: PF00581; Rhodanese; 2.  
 DR SMART: SM00450; RHOD; 2.  
 DR PROSITE: PS00360; RHODANESE\_1; 1.  
 DR PROSITE: PS00683; RHODANESE\_2; 1.  
 DR Transferrase; Mitochondrion.  
 FT DOMAIN 1 142  
 FT DOMAIN 143 158  
 FT HINGE.

FT DOMAIN 159 289  
 FT ACT\_SITE 186 186  
 FT ACT\_SITE 244 244  
 FT ACT\_SITE 245 245  
 FT ACT\_SITE 246 246  
 FT ACT\_SITE 246 246  
 FT SEQUENCE 289 AA; 32286 MW; 8BFCF671DE0B2BA4 CRC64;

Query Match 76.3%; Score 29; DB 1; Length 289;

Best Local Similarity 83.3%; Pred. No. 42;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 2 RPVGRF 7  
 || ||  
 Db 182 RPACRF 187

RESULT 13

AR72\_HUMAN STANDARD; PRT; 330 AA.

ID AR72\_HUMAN

AC O43458; O25749;

DT 15-JUL-1999 (Rel. 38, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 16-OCT-2001 (Rel. 40, Last annotation update)

DE Aflatoxin B1 aldehyde reductase 1 (EC 1.-.-.-) (AFB1-AR 1)

DE (Aldoketoreductase 7).

GN AKR7A2 OR AFAR OR AKR7.

OS Homo sapiens (Human).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.

OC NCBI\_TaxID=9606;

RN [1]

RP SEQUENCE FROM N.A.

RC TISSUE=Liver;

RA MEDLINE=98244807; PubMed=9576847;

RT Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;

RT "Molecular cloning, expression and catalytic activity of a human AKR7

RT member of the aldo-keto reductase superfamily: evidence that the

RT major 2-carboxybenzaldehyde reductase from human liver is a homologue

RT of rat aflatoxin B1-aldehyde reductase.";

RL Biochem. J. 332:21-34(1998).

RN [2]

RP SEQUENCE FROM N.A.

RC TISSUE=Brain;

RA MEDLINE=99040634; PubMed=9823300;

RT Praml C., Savelyeva L., Perri P., Schwab M.;

RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-

RT 1p36.1 in a region frequently altered in human tumor cells.";

RL Cancer Res. 58:5014-5018(1998).

RN [3]

RP SEQUENCE FROM N.A.

RA Hall R.;

RT Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.

CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING

CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOL. COULD BE

CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC

CC EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE

CC REDUCTASE.

CC -1- SUBCELLULAR LOCATION: Cytoplasmic.

CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: AF026947; AAC52104.1; -.



```

DR EMBL; Y16675; CAA76347.1; -.
DR EMBL; AL035413; CAB72321.1; -.
DR MIM; 603418; -.
DR InterPro; IPR001395; Aldo_ket_red.
DR Pfam; PF00248; aldo_ket_red; 1.
KW Oxidoreductase.
FT ACT_SITE 112 112 HYDROGEN-BOND DONOR (PROBABLE).
FT CONFLICT 113 113 A -> T (IN REF. 1).
SQ SEQUENCE 330 AA; 36618 MM; 3BBFB7ED0CAFA4D54 CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 330;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPVGRF 7
DB 217 QPVGRF 222

RESULT 14
AR73_HUMAN
ID AR73_HUMAN STANDARD: PRT: 331 AA.
AC 095154; Q9NUG3;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Aflatoxin B1 aldehyde reductase 2 (EC 1.-.-.-) (AFB1-AR 2).
GN AKR7A3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
[1]
RP SEQUENCE FROM N.A. AND CHARACTERIZATION.
RC TISSUE=Liver;
RX MEDLINE=99315412; Pubmed=10383892;
RA Knight L.P., Primiano T., Groopman J.D., Kensler T.W., Sutter T.R.;
RT "cDNA cloning, expression and activity of a second human aflatoxin
RT B1-metabolizing member of the aldo-keto reductase superfamily,
RT AKR7A3."
RL Carcinogenesis 20:1215-1223(1999).
RN
RP SEQUENCE FROM N.A.
RA Hall R.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC
CC EFFECTS OF AFLATOXIN B1.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AF040639; AAD02195.1; -.
CC EMBL; AL035413; CAB72322.1; -.
CC InterPro; IPR001395; Aldo_ket_red.
CC Pfam; PF00248; aldo_ket_red; 1.
CC PRINTS; PR00069; ALDKETREDTASE.
KW Oxidoreductase.
FT ACT_SITE 113 113 HYDROGEN-BOND DONOR (PROBABLE).
FT CONFLICT 51 51 E -> D (IN REF. 1).
FT CONFLICT 138 138 V -> M (IN REF. 1).
FT CONFLICT 201 201 A -> ADGSPGCGSFMCTLGRGADCCFPS (IN REF.
FT CONFLICT 215 215 N -> D (IN REF. 1).

```

```

FT CONFLICT 323 323 T -> A (IN REF. 1).
SQ SEQUENCE 331 AA; 37206 MM; B9C32C3C3C7102AB3 CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 331;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPVGRF 7
DB 218 QPVGRF 223

RESULT 15
RPA2_PYRAB
ID RPA2_PYRAB STANDARD: PRT: 397 AA.
AC 09V113;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE DNA-directed RNA polymerase subunit A" (EC 2.7.7.6).
GN RPOA2 OR PAB0425.
GN Pyrococcus abyssi.
OS Archaea; Euryarchaeota; Thermococcales; Thermococcaceae; Pyrococcus.
OX NCBI_TaxID=29292;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=ORSAY;
RA Helling R.;
RT "Pyrococcus abyssi genome sequence: insights into archaeal chromosome
RT structure and evolution."
RL Submitted (JUL-1999) to the EMBL/GenBank/DBJ databases.
CC -1- FUNCTION: DNA-DEPENDENT RNA POLYMERASE CATALYZES THE TRANSCRIPTION
CC OF DNA INTO RNA USING THE FOUR RIBONUCLEOSIDE TRIPHOSPHATES AS
CC SUBSTRATES.
CC -1- CATALYTIC ACTIVITY: N nucleoside triphosphate - N diphosphate +
CC {RNA}(N).
CC -1- SIMILARITY: BELONGS TO THE RNA POLYMERASE BETA' CHAIN FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
CC EMBL; AJ248284; CAB49538.1; -.
CC InterPro; IPR002879; RNA_pol_A2.
CC Pfam; PF01854; RNA_pol_A2; 1.
CC Transferrase; Transcription; DNA-directed RNA polymerase;
KW Complete proteome.
SQ SEQUENCE 397 AA; 44594 MM; EA59658EAB9C15CB CRC64;

Query Match
Best Local Similarity 76.3%; Score 29; DB 1; Length 397;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVG 6
DB 320 VRPGR 325

```

Search completed: September 13, 2002, 09:30:49  
Job time: 1139 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:25 ; Search time 311.85 seconds

(without alignments)  
4.438 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_14\_21

Perfect score: 38

Sequence: 1 IRPVGRFX 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: SP archaea: \*  
2: SP bacteria: \*  
3: SP fungi: \*  
4: SP human: \*  
5: SP invertebrate: \*  
6: SP mammal: \*  
7: SP mnc: \*  
8: SP organelle: \*  
9: SP phage: \*  
10: SP plant: \*  
11: SP rodent: \*  
12: SP virus: \*  
13: SP vertebrate: \*  
14: SP unclassified: \*  
15: SP viirus: \*  
16: SP bacteriaph: \*  
17: SP archaeap: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	37	97.4	420	2	Q931Z7 streptomyc
2	35	92.1	117	13	Q9W6Z4
3	34	89.5	529	3	Q9C1M8
4	33	86.8	445	2	Q50531
5	33	86.8	511	3	Q59714
6	31	81.6	612	2	Q9X8C4
7	30	78.9	154	2	Q54139
8	30	78.9	173	16	P74584
9	30	78.9	198	10	Q9C7Y9
10	30	78.9	274	10	Q9SA12
11	30	78.9	292	4	Q9H6Y1
12	30	78.9	292	11	Q9D8X5
13	30	78.9	313	4	Q9NXM2
14	30	78.9	313	5	Q9U8B2
15	30	78.9	318	10	Q9AVS9
16	30	78.9	369	5	Q961L5

17	30	78.9	373	10	Q9SUS8	Q9SUS8 arabidopsis
18	30	78.9	374	10	Q41745	Q41745 zea mays (m
19	30	78.9	393	10	Q9M1Z5	Q9M1Z5 arabidopsis
20	30	78.9	406	10	Q82361	Q82361 arabidopsis
21	30	78.9	426	17	Q979S0	Q979S0 thermoplas
22	30	78.9	435	16	Q910E2	Q910E2 pseudomon
23	30	78.9	438	4	Q9NVA4	Q9NVA4 homo sapien
24	30	78.9	501	3	Q08229	Q08229 saccharomyc
25	30	78.9	533	5	Q18684	Q18684 caenorhabdi
26	30	78.9	559	16	Q9KBES	Q9KBES bacillus ha
27	30	78.9	572	11	Q70490	Q70490 rattus norv
28	30	78.9	601	17	Q300J9	Q300J9 archaeglob
29	30	78.9	609	16	Q99RJ1	Q99RJ1 staphylococ
30	30	78.9	622	11	Q99KQ7	Q99KQ7 mus musculu
31	30	78.9	674	16	Q50431	Q50431 mycobacteri
32	30	78.9	751	16	Q9HVO1	Q9HVO1 pseudomonas
33	30	78.9	2546	2	Q9A130	Q9A130 burkholderi
34	29	76.3	90	10	Q9S809	Q9S809 oryza longi
35	29	76.3	91	10	Q9S7S0	Q9S7S0 oryza sativ
36	29	76.3	91	10	Q9ZMW5	Q9ZMW5 oryza longi
37	29	76.3	95	12	Q9WC44	Q9WC44 hepatitis e
38	29	76.3	95	12	Q9WC46	Q9WC46 hepatitis e
39	29	76.3	95	12	Q9WC47	Q9WC47 hepatitis e
40	29	76.3	95	12	Q9WC51	Q9WC51 hepatitis e
41	29	76.3	95	12	Q912D6	Q912D6 hepatitis e
42	29	76.3	104	8	Q34311	Q34311 dictyostell
43	29	76.3	109	13	Q90Y59	Q90Y59 paraliichth
44	29	76.3	114	12	Q91PL7	Q91PL7 hepatitis c
45	29	76.3	118	4	Q9P195	Q9P195 homo sapien

## ALIGNMENTS

RESULT 1  
Q931Z7 PRELIMINARY; PRT; 420 AA.  
AC Q931Z7;  
DT 01-DEC-2001 (TREMBLrel. 19, Created)  
DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CHAIN LENGTH FACTOR-LIKE PROTEIN.  
GN AUR2B.  
OS Streptomyces aureofaciens.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1894;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=CCM3239;  
RA Kormanec J., Bistakova J., Novakova R., Homerova D., Rezuchova B.;  
RT "Cloning and characterization of a new polyketide gene cluster in  
RT Streptomyces aureofaciens CCM3239.";  
RL Submitted (May-2001) to the EMBL/Genbank/DBJ databases.  
DR EMBL; AY033994; AAK61719.1; -  
SQ SEQUENCE 420 AA; 43011 MW; 3C27E22BE88C2DEA CRC64;

Query Match 97.4%; Score 37; DB 2; Length 420;  
Best Local Similarity 100.0%; Pred. No. 6.5;  
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
|||||  
DB 49 IRPVGRF 55

RESULT 2  
Q9W6Z4 PRELIMINARY; PRT; 117 AA.  
AC Q9W6Z4;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)

DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)  
DE C-RF AMIDE.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Salake H., Minakata H., Fujimoto M.;  
RT "Carassius Rhamde (C-RF amide).";  
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AB020024; BAF76662.1; -  
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB22038C2B0 CRC64;

Query Match 92.1%; Score 35; DB 13; Length 117;  
Best Local Similarity 71.4%; Pred. NO. 4.9;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IRPYGRF 7  
:|||||  
Db 69 VRPIGRF 75

RESULT 3  
09C1M8 PRELIMINARY; PRT; 529 AA.  
ID 09C1M8  
AC 09C1M8;  
DT 01-JUN-2001 (TEMBLrel. 17, Created)  
DT 01-JUN-2001 (TEMBLrel. 17, last sequence update)  
DT 01-DEC-2001 (TEMBLrel. 19, last annotation update)  
DE HEAT-INDUCED CATALASE.  
OC Pleurotus sajor-caju (Oyster mushroom).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
OC Agaricales; Pleurotaceae; Pleurotus.  
OX NCBI\_TaxID=50053;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Jeong M.-J., Park S.-C.;  
RT "Cloning of the catalase gene from Pleurotus sajor-caju.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AP286097; AKI5159.1; -  
DR HSSP: P15202; 1A4E.  
DR Interpro: IPR002226; Catalase.  
DR Pfam: PF00199; catalase.1.  
DR PRINTS: PR00067; CATALASE.  
DR ProDom: PD000510; CATALASE.  
DR PROSITE: PS00437; CATALASE\_1.1.  
DR PROSITE: PS00438; CATALASE\_2; UNKNOWN\_1.  
SQ SEQUENCE 529 AA; 59791 MW; A432BE6445B6AC13 CRC64;

Query Match 89.5%; Score 34; DB 3; Length 529;  
Best Local Similarity 71.4%; Pred. NO. 37;  
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 IRPYGRF 7  
:|||||  
Db 298 LRPIGRF 304

RESULT 4  
050531 PRELIMINARY; PRT; 445 AA.  
ID 050531  
AC 050531;  
DT 01-JUN-1998 (TEMBLrel. 06, Created)  
DT 01-JUN-1998 (TEMBLrel. 06, last sequence update)  
DT 01-JUN-2001 (TEMBLrel. 17, last annotation update)  
DE FAD-DEPENDENT OXIDOREDUCTASE.  
GN SC9810.24C.  
OS Streptomyces coelicolor.

OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-A3(2).  
RA Oliver K., Harris D.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-A3(2);  
RA Parthill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-A3(2);  
RX MEDLINE=97000351; Pubmed=8443436;  
RA Redenbach M., Kleser H.M., Denapate D., Eichner A., Cullum J.,  
RA Kinashi H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL MOL. Microbiol. 21:77-96(1996).  
DR EMBL: AL009204; CAA15814.1; -  
DR Interpro: IPR001575; Oxid\_FAD\_bind.  
DR Pfam: PF01565; FAD\_binding\_4.1.  
DR PROSITE: PS00862; OX2\_COVAL\_FAD; UNKNOWN\_1.  
SQ SEQUENCE 445 AA; 49067 MW; 8C0AEFB874E38CF5 CRC64;

Query Match 86.8%; Score 33; DB 2; Length 445;  
Best Local Similarity 100.0%; Pred. NO. 51;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 2 RPYGRF 7  
|||||  
Db 248 RPYGRF 253

RESULT 5  
059714 PRELIMINARY; PRT; 511 AA.  
ID 059714  
AC 059714;  
DT 01-AUG-1998 (TEMBLrel. 07, Created)  
DT 01-AUG-1998 (TEMBLrel. 07, last sequence update)  
DT 01-MAR-2001 (TEMBLrel. 16, last annotation update)  
DE HYPOTHETICAL 57.7 KDA PROTEIN C3B8.06 IN CHROMOSOME II.  
GN SPBC3B8.06.  
OS Schizosaccharomyces pombe (Fission Yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC SRRAIN-972;  
RA Lyne M., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.,  
RA Pohl T.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: IN C-TERMINAL TO YEAST YCR062W.  
DR EMBL: AL022244; CAA18295.1; -  
DR Hypothetical protein; Transmembrane.  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 33 53 POTENTIAL.  
FT TRANSMEM 59 79 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 332 352 POTENTIAL.  
FT TRANSMEM 379 399 POTENTIAL.  
FT DOMAIN 419 424 POLY-ALA.  
FT TRANSMEM 449 469 POTENTIAL.  
FT TRANSMEM 483 503 POTENTIAL.  
SQ SEQUENCE 511 AA; 57703 MW; BDDDE7650B2225E CRC64;

Query Match 86.8%; Score 33; DB 3; Length 511;  
Best Local Similarity 100.0%; Pred. No. 58;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPYGRF 7  
| | | | |  
DB 95 RPYGRF 100

RESULT 6  
09X8C4 PRELIMINARY; PRT: 612 AA.  
AC 09X8C4;  
DT 01-NOV-1999 (TREMBLrel. 12, Created)  
DT 01-NOV-1999 (TREMBLrel. 12, Last sequence update)  
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)  
DE CONSERVED HYPOTHETICAL PROTEIN (FRAGMENT).  
GN SC236.11C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Oliver K., Harris D.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Bentley S.D., Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC MEDLINE-97000351; PubMed-8843436;  
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
RA Kinasht H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL: AL049763; CABA2080.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 612 AA; 66978 MW; 6CDECEB345D26D3 CRC64;

Query Match 81.6%; Score 31; DB 2; Length 612;  
Best Local Similarity 85.7%; Pred. No. 1.9e+02;  
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPYGRF 7  
| | | | |  
DB 49 ITPVGRF 55

RESULT 7  
054139 PRELIMINARY; PRT: 154 AA.  
AC 054139;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)  
DT 01-NOV-1998 (TREMBLrel. 08, Last annotation update)  
DE HYPOTHETICAL 16.9 KDA PROTEIN.  
GN SC239.14.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Oliver K., Harris D.;

RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Parkhill J., Barrell B.G., Rajandream M.A.;  
RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RX MEDLINE-97000351; PubMed-8843436;  
RA Redenbach M., Kieser H.M., Denapalte D., Eichner A., Cullum J.,  
RA Kinasht H., Hopwood D.A.;  
RT "A set of ordered cosmids and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL: AL021530; CAA16482.1; -.  
KW Hypothetical protein.  
FT NON\_TER 1  
SQ SEQUENCE 154 AA; 16884 MW; 9D9BA7F36B84CD03 CRC64;

Query Match 78.9%; Score 30; DB 2; Length 154;  
Best Local Similarity 83.3%; Pred. No. 79;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPYGRF 7  
: | | | | |  
DB 31 RPYGRF 36

RESULT 8  
P74584 PRELIMINARY; PRT: 173 AA.  
ID P74584;  
AC P74584;  
DT 01-FEB-1997 (TREMBLrel. 02, Created)  
DT 01-FEB-1997 (TREMBLrel. 02, Last sequence update)  
DT 01-JUN-2000 (TREMBLrel. 14, Last annotation update)  
DE HYPOTHETICAL 20.5 KDA PROTEIN.  
GN SLR0667.  
OS Synecocystis sp. (strain PCC 6803).  
OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
OX NCBI\_TaxID=1148;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE-97061201; PubMed-8905231;  
RA Kaneke T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
RA Miyajima N., Hiroseawa M., Sugiyama M., Sasamoto S., Kimura T.,  
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
RA Shimpou S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
RA Tabata S.;  
RT "Sequence analysis of the genome of the unicellular cyanobacterium  
Synecocystis sp. strain PCC6803. II. Sequence determination of the  
entire genome and assignment of potential protein-coding regions.";  
RL DNA Res. 3:109-136(1996).  
DR EMBL: D90916; BAA18691.1; -.  
KW Hypothetical protein; Complete proteome.  
FT NON\_TER 1  
SQ SEQUENCE 173 AA; 20509 MW; 2E2414F099C8B2F7 CRC64;

Query Match 78.9%; Score 30; DB 16; Length 173;  
Best Local Similarity 83.3%; Pred. No. 88;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPYGRF 7  
: | | | | |  
DB 94 RPYGRF 99

RESULT 9  
09C7Y9 PRELIMINARY; PRT: 198 AA.  
ID 09C7Y9;  
AC 09C7Y9;  
DT 01-JUN-2001 (TREMBLrel. 17, Created)  
DT 01-JUN-2001 (TREMBLrel. 17, Last sequence update)

```

DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE HYPOHETICAL 21.8 KDA PROTEIN (UNKNOWN PROTEIN).
GN T2J15.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RX MEDLINE=21016719; PubMed=11130712;
RA Theologis A., Ecker J.R., Palm C.J., Federspiel N.A., Kaul S.,
RA White O., Alonso J., Altati H., Araujo R., Bowmen C.L., Brooks S.Y.,
RA Buehler E., Chao A., Chen Q., Chen H., Cheuk R.F., Chin C.W.,
RA Chung M.K., Conn L., Conway A.B., Conway A.R., Creasy T.H., Dewar K.,
RA Dunn P., Elgu P., Feldblum T.V., Feng J.-D., Fong B., Fujii C.Y.,
RA Gill J.E., Goldsmith A.D., Haas B., Hansen N.F., Hughes B., Huzar L.,
RA Hunter J.L., Jenkins J., Johnson-Hopson C., Khan S., Khaykin E.,
RA Kim C.J., Koo H.L., Kremetska I., Kurtz D.B., Kwan A., Lam B.,
RA Langin-Hooper S., Lee A., Lee J.M., Lenz C.A., Li J.H., Li Y.-P.,
RA Lin X., Liu S.X., Liu Z.A., Lucos J.S., Malt R., Marziani A.,
RA Miltescher J., Miranda M., Nguyen M., Nierman W.C., Osborne B.I.,
RA Pal G., Peterson J., Pham P.K., Rizzo M., Rooney T., Rowley D.,
RA Sakano H., Salzberg S.L., Schwartz J.R., Shinn P., Southwick A.M.,
RA Sun H., Tallon L.J., Tambunga G., Tortum M.J., Town C.D.,
RA Ulteback T., Van Aken S., Vaysberg M., Vysotskaya V.S., Walker M.,
RA Wu D., Yu G., Fraser C.M., Venter J.C., Davis R.W.;
RT *Sequence and analysis of chromosome 1 of the plant Arabidopsis
RT thaliana.
RT Nature 408:816-820(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Lam B., Southwick A., Karlin-Neumann G., Nguyen M., Miranda M.,
RA Palm C.J., Bower L., Jones T., Banh J., Carlini P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kaniya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shin P., Yamada K., Shinozaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC051631; AAG51532.1; -.
DR EMBL; AF386997; AAK62442.1; -.
DR Hypothetical protein
SQ SEQUENCE 198 AA; 21847 MW; C0AC434A40FE7818 CRC64;

Query Match
Best Local Similarity 83.3%; Score 30; DB 10; Length 198;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGR 6
DB 127 VRPVGR 132

RESULT 10
O9SAI2 PRELIMINARY; PRT; 274 AA.
AC O9SAI2;
DT 01-MAY-2000 (TEMBLrel. 13, Created)
DT 01-MAY-2000 (TEMBLrel. 13, Last sequence update)
DT 01-DEC-2001 (TEMBLrel. 19, Last annotation update)
DE F23A5.13 PROTEIN.
GN Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Schwartz J.R., Yu G., Tortum M., Lenz C., Liu S., Lee J.M., Li J.,

```

```

RA Gonzalez A., Liu A., Liu K., Sakano H., Koo T., Pham P., Vaysberg M.,
RA Howing B., Chin C., Choi E., Chlou J., Altafi H., Brooks S., Chao Q.,
RA Conn L., Conway A., Hansen N., Johnson-Hopson C., Khan S., Kim C.,
RA Lam B., Nguyen M., Palm C., Shinn P., Tambunga G., Davis R.W.,
RA Ecker J.R., Federspiel N.A., Theologis A.,
RT *Arabidopsis thaliana chromosome 1 BAC F23A5 sequence.
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CV, COLUMBIA;
RA Theologis A.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC011713; AAF14666.1; -.
SQ SEQUENCE 274 AA; 31200 MW; 6D595880036C9CC CRC64;

Query Match
Best Local Similarity 71.4%; Score 30; DB 10; Length 274;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGR 7
DB 48 VRPVGR 54

RESULT 11
O9H6Y1 PRELIMINARY; PRT; 292 AA.
ID O9H6Y1;
AC O9H6Y1;
DT 01-MAR-2001 (TEMBLrel. 16, Created)
DT 01-MAR-2001 (TEMBLrel. 16, Last sequence update)
DE CDNA: FLJ21705 FIS, CLONE COL09940.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=COLON;
RA Kawabata A., Hiki J., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Ohtani R., Ota T., Suzuki Y., Ohyashi M., Mishi T., Shiohara T.,
RA Tanaka T., Nakamura Y., Isogai T., Sugano S.;
RT *NEO human cDNA sequencing project.
RL Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AK025358; BAB15119.1; -.
SQ SEQUENCE 292 AA; 33568 MW; 8102682193DB40E6 CRC64;

Query Match
Best Local Similarity 57.1%; Score 30; DB 4; Length 292;
Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGR 7
DB 48 VRPIGR 54

RESULT 12
O9D8X5 PRELIMINARY; PRT; 292 AA.
ID O9D8X5;
AC O9D8X5;
DT 01-JUN-2001 (TEMBLrel. 17, Created)
DT 01-JUN-2001 (TEMBLrel. 17, Last sequence update)
DT 01-JUN-2001 (TEMBLrel. 17, Last annotation update)
DE 1810022F04RIK PROTEIN (RIKEN CDNA 1810022F04 GENE).
GN 1810022F04RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.

```

RC STRAIN=C57BL/6J; TISSUE=PANCREAS;  
 RA MEDLINE=21085660; Pubmed=11217851;  
 RA Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,  
 RA Arikawa T., Hara A., Fukunishi Y., Kono H., Adachi J., Fukuda S.,  
 RA Aizawa K., Iwata M., Nishi K., Kiyosawa H., Kondo S., Yamanka I.,  
 RA Salto T., Okazaki Y., Gojohori T., Bono H., Kasukawa T., Saito R.,  
 RA Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,  
 RA Flaischmann W., Gaasterland T., Gissi C., King B., Kochwa H.,  
 RA Kuehl P., Lewis S., Matsuo Y., Nakai I., Pesole G., Quackenbush J.,  
 RA Schriml L.M., Staahl F., Suzuki R., Tomita M., Wagner L., Washio T.,  
 RA Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,  
 RA Blake J., Bottelli D., Bojunga N., Carninci P., de Bonaldo M.F.,  
 RA Brownstein M.J., Bult C., Fletcher C., Fujita M., Gariboldi M.,  
 RA Guenlinch S., Hill D., Holmann M., Hume D.A., Kamita M., Lee N.H.,  
 RA Lyons P., Marchionni L., Mashima J., Mazarelli J., Mombarts P.,  
 RA Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,  
 RA Sasaki H., Sato K., Schoenbach C., Seiya T., Shibata Y., Storch K.-F.,  
 RA Suzuki H., Toyooka K., Wang K.H., Wetz C., Whitaker C., Wilmink L.,  
 RA Wyshew-Boris A., Yoshida K., Haegawa Y., Kawai H., Kohlschki S.,  
 RA Hayashizaki Y.,  
 RT \*Functional annotation of a full-length mouse cDNA collection.\*;  
 RL Nature 409:685-690(2001).  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK007581; BAB5119.1; -  
 DR EMBL: BC004040; AAH04040.1; -  
 DR MGI: MGI:1916375; 1810022F04RLK.  
 SO SEQUENCE 292 AA; 33574 MW; 9CD57A966E51A91B CRC64;

Query Match 78.9%; Score 30; DB 11; Length 292;  
 Best Local Similarity 57.1%; Pred. No. 1.5e+02;  
 Matches 4; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
 :||:|  
 DB 48 VARIPEF 54

RESULT 13  
 ID 09NKM2 PRELIMINARY; PRT; 313 AA.  
 AC 09NKM2;  
 DT 01-OCT-2000 (TREMBlrel. 15, Created)  
 DT 01-OCT-2000 (TREMBlrel. 15, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE CDNA FLJ20162 FIS, CLONE COL09280.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.  
 OC NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=COLON;  
 RA Kawabata A., Hiki J., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,  
 RA Okitani R., Ota T., Suzuki Y., Obayashi M., Nishi T., Shibahara T.,  
 RA Tanaka T., Nakamura Y., Isogai T., Sugano S.,  
 RT \*NEDO human cDNA sequencing project.\*;  
 RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AK000169; BAA90988.1; -  
 SO SEQUENCE 313 AA; 35351 MW; B3AE53EAB738869 CRC64;

Query Match 78.9%; Score 30; DB 4; Length 313;  
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGRF 7  
 :||:|  
 DB 117 IOPVGRF 123

RESULT 14  
 ID 09UB82 PRELIMINARY; PRT; 313 AA.  
 AC 09UB82;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DE CARBAMATE KINASE (EC 2.7.2.2).  
 GN CBK.  
 OS Hexamita sp.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Hexamitinae; Hexamita.  
 OC NCBI\_TaxID=28003;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA MEDLINE=20454785; Pubmed=11001147;  
 RA Dimopoulos M., Bagnara A.S., Edwards M.R.;  
 RT "Characterisation and sequence analysis of a carbamate kinase gene  
 from the diplomonad Hexamita inflata.";  
 RT J. Eukaryot. Microbiol. 47:499-503(2000).  
 RL EMBL: AF107491; AAF08984.1; -  
 DR HSSP: P95474; 1E19.  
 DR InterPro: IPR001048; Aakkinase.  
 DR InterPro: IPR003964; Bact\_carbamate\_kinase.  
 DR InterPro: IPR001057; Glut\_5\_kinase.  
 DR Pfam: PF00696; aakkinase; 1.  
 DR PRINTS: PR00474; GLU5KINASE.  
 KW Kinase; Transferease.  
 SO SEQUENCE 313 AA; 33739 MW; 52D3AB67B7E19278 CRC64;

Query Match 78.9%; Score 30; DB 5; Length 313;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 2 RPYGRF 7  
 :||||  
 DB 130 KPYGRF 135

RESULT 15  
 ID 09AV59 PRELIMINARY; PRT; 318 AA.  
 AC 09AV59;  
 DT 01-JUN-2001 (TREMBlrel. 17, Created)  
 DT 01-JUN-2001 (TREMBlrel. 17, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE HYPOTHETICAL 33.7 KDA PROTEIN.  
 GN OSJNB0093B11.17.  
 OS Oryza sativa (Rice).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Euphorbioidae; Oryzae; Oryza.  
 OC NCBI\_TaxID=4530;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=CV. NIPONBARE;  
 RA Buell C.R., Yuan Q., Ouyang S., Moffat K.S., Hill J.N., Burr P.C.,  
 RA Hsiao J., Zismann V., Pal G., Bowman C.L., Fujii C.Y., Vanaken S.E.,  
 RA Bowman G.L., Craven B., Ulterback T.R., Khalak H., Feldblum T.V.,  
 RA Quackenbush J., White O., Salzberg S.L., Fraser C.M.;  
 RT "Oryza sativa chromosome 10 BAC OSJNB0093B11 genomic sequence.";  
 RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: AC024594; AAK21337.1; -  
 KW Hypothetical protein.  
 SO SEQUENCE 318 AA; 33656 MW; A92434254D3F7D93 CRC64;

Query Match 78.9%; Score 30; DB 10; Length 318;  
 Best Local Similarity 83.3%; Pred. No. 1.6e+02;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 IRPVGR 6

Db :|||||  
20 VPRVGR 25

Search completed: September 13, 2002, 09:29:26  
Job time: 1071 sec

---



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:18:36 ; Search time 399.68 Seconds  
(without alignments)  
1.945 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_15\_21

Perfect score: 34

Sequence: 1 RPVGPRX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 747574 seqs, 111073796 residues

Total number of hits satisfying chosen parameters: 747574

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A\_Geneseq\_032802:\*

1: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1980.DAT:\*  
2: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1981.DAT:\*  
3: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1982.DAT:\*  
4: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1983.DAT:\*  
5: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1984.DAT:\*  
6: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1985.DAT:\*  
7: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1986.DAT:\*  
8: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1987.DAT:\*  
9: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1988.DAT:\*  
10: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1989.DAT:\*  
11: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1990.DAT:\*  
12: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1991.DAT:\*  
13: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1992.DAT:\*  
14: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1993.DAT:\*  
15: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1994.DAT:\*  
16: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1995.DAT:\*  
17: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1996.DAT:\*  
18: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1997.DAT:\*  
19: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1998.DAT:\*  
20: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA1999.DAT:\*  
21: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA2000.DAT:\*  
22: /SIDSI/gcgdata/hold-genseq/genseqp-emb1/AA2001.DAT:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	97.1	8	22	AA046955
2	33	97.1	10	20	AA095178
3	33	97.1	15	18	AA031400
4	33	97.1	15	20	AA097230
5	33	97.1	15	21	AA049293
6	33	97.1	15	21	AA049295
7	33	97.1	15	21	AA049296
8	33	97.1	20	18	AA031394
9	33	97.1	20	18	AA031387
10	33	97.1	20	18	AA031374
11	33	97.1	20	20	AA097232

12	33	97.1	20	20	AA097234	Rat type ligand po
13	33	97.1	20	20	AA097236	Human type ligand
14	33	97.1	20	20	AA095191	Bovine pituitary-d
15	33	97.1	20	20	AA095175	Murine pituitary-d
16	33	97.1	20	21	AA010350	Bovine oxytocin se
17	33	97.1	20	21	AA010358	Rat oxytocin secre
18	33	97.1	20	21	AA010365	Human oxytocin sec
19	33	97.1	20	21	AA010369	Oxytocin secretion
20	33	97.1	20	21	AA049294	19P2 ligand peptid
21	33	97.1	20	21	AA049301	19P2 ligand peptid
22	33	97.1	20	21	AA049302	19P2 ligand peptid
23	33	97.1	20	22	AA062519	Bovine CRH releas
24	33	97.1	20	22	AA062527	Rat CRH releas
25	33	97.1	20	22	AA062534	Human CRH releas
26	33	97.1	20	22	AA062538	CRH releas
27	33	97.1	20	22	AA090992	Proactin releas
28	33	97.1	20	22	AA090994	Proactin releas
29	33	97.1	20	22	AA090996	Proactin releas
30	33	97.1	20	22	AA046954	Synthetic G prote
31	33	97.1	20	18	AA031397	Synthetic G prote
32	33	97.1	21	18	AA031395	Human type G prote
33	33	97.1	21	18	AA031388	Rat type G prote
34	33	97.1	21	18	AA031375	Bovine G protein-c
35	33	97.1	21	20	AA097227	Partial ligand pol
36	33	97.1	21	20	AA067616	Mammalian 19P2 lig
37	33	97.1	21	20	AA095191	Bovine pituitary-d
38	33	97.1	21	21	AA010351	Bovine oxytocin se
39	33	97.1	21	21	AA010359	Rat oxytocin secre
40	33	97.1	21	21	AA010366	Human oxytocin sec
41	33	97.1	21	22	AA062520	Bovine CRH releas
42	33	97.1	21	22	AA062528	Rat CRH releas
43	33	97.1	21	22	AA062535	Human CRH releas
44	33	97.1	22	18	AA031396	Human type G prote
45	33	97.1	22	18	AA031389	Rat type G prote

#### ALIGNMENTS

RESULT	1
AA046955	
ID	AA046955 standard; Protein; 8 AA.
XX	
AC	AA046955;
XX	
DT	04-MAY-2001 (first entry)
XX	
DE	Peptide PRP8 fragment.
XX	
KW	GPR10; UHR-1; PRP receptor; prolactin-releasing peptide; pain;
KW	central nervous system disorder; autonomic regulation; analgesic;
KW	hypotensive; blood pressure.
XX	
OS	Unidentified.
XX	
FT	Key
FT	Modified-site
FT	8
XX	/note="C-terminal amide"
XX	
PN	WO200109182-A1.
XX	
PD	08-FEB-2001.
XX	
PF	03-AUG-2000; 2000WO-F100664.
XX	
PR	03-AUG-1999; 99US-0365756.
XX	
PA	20-MAR-2000; 2000US-031367.
XX	
PI	(JUVA-) JUVANTIA PHARMA LTD OY.
XX	
DR	Panula PAJ, Pertovaara A, Kalso E, Korpi E;
XX	
DR	WPI; 2001-182941/18.

XX C-terminal fragments of prolactin-releasing peptide useful for  
PT regulating autonomic functions and in the manufacture of a medicament  
PT for regulating blood pressure  
PS Claim 2: Page 10; 40pp; English.

XX This invention describes a novel C-terminal fragment (I) of an isolated  
CC prolactin-releasing peptide (PrRP), referred to as PrRP20 and having a  
CC sequence (S1). The invention also describes (1) a therapeutic composition  
CC (C1) comprising (I) or a C-terminal fragment of PrRP referred to as PrRP8  
CC and comprising a sequence (S2); (2) a diagnostic method based on antisera  
CC against PrRP20 for identification of disorders involving the central  
CC nervous system, including those associated with pain or autonomic  
CC regulation, where specific antisera against the N-and/or C-terminal  
CC domains of PrRP is used to identify alterations in PrRP synthesis or  
CC levels; (3) a rat or human receptor encoded by a 1122 nucleotide sequence  
CC (S3), fully defined in the specification; (4) treating a person suffering  
CC from a disorder regulated by a receptor (II) encoded by a sequence of  
CC 1122 nucleotide sequence, fully defined in the specification, located in  
CC the central nervous system, by administering an agonist or antagonist to  
CC the receptor; and (5) treating blood pressure, by blocking of receptors  
CC of PrRP or its C-terminal fragment GlyIleArgProValGlyArgphe-NH<sub>2</sub> (S7).  
CC The products of the invention have analgesic and hypotensive activity.  
CC (I) is useful for regulating autonomic functions, such as increasing  
CC blood pressure. (1) is useful for treating pain, for manufacturing a  
CC medicament for regulating blood pressure, and for treating pain. Agonist  
CC and antagonist of (II) are useful for treating acute pain, inflammatory  
CC pain and neuropathic pain, for regulating autonomic functions and  
CC treating high blood pressure.

XX Sequence 8 AA:  
SQ

Query Match 97.1%; Score 33; DB 22; Length 8;  
Best Local Similarity 100.0%; Pred. No. 6.4e+05;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
| | | | |  
DB 3 rpygrf 8

RESULT 2  
AAW95178  
ID AAW95178 standard; Protein; 10 AA.  
XX  
AC AAW95178;  
XX  
DT 10-MAR-1999 (first entry)  
XX  
DE Murine pituitary-derived ligand polypeptide antigenic epitope.  
XX  
KW Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
KM GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KW tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KW Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KW Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KW secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KW gene therapy; transgenic animal; epitope.  
XX  
OS Mus sp.  
XX  
PN WO9849295-A1.  
XX  
PD 05-NOV-1998.  
XX  
PF 27-APR-1998; 98WO-JP01923.  
XX  
PR 28-APR-1997; 97JP-0109974.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Fukusumi S, Hinuma S;  
XX  
DR WPI; 1999-009423/01.  
XX  
PT New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening  
PS Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutelin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.  
XX

XX Sequence 10 AA:  
SQ

Query Match 97.1%; Score 33; DB 20; Length 10;  
Best Local Similarity 100.0%; Pred. No. 0.92;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
| | | | |  
DB 5 rpygrf 10

RESULT 3  
AAW31400  
ID AAW31400 standard; Peptide; 15 AA.  
XX  
AC AAW31400;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Synthetic ligand 19P2-L31 peptide II.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KM modulator; pituitary; central nervous system; pancreas; propylactic;  
KW therapeutic agent; antigen.  
XX  
OS Synthetic.  
XX  
PN WO9724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PF 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
XX  
PR 28-DEC-1995; 95JP-0343371.  
XX  
PR 15-MAR-1996; 96JP-0059419.  
XX  
PR 12-AUG-1996; 96JP-0211805.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
XX Fuji R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kltada C;  
XX  
XX WPI: 1997-363672/33.  
XX  
XX Lligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
XX  
XX Example 43; Page 151; 258pp; English.  
XX  
XX This peptide contains the partial C-terminal sequence of the synthetic  
CC ligand polypeptide 19p2-L31 which is capable of binding to a G  
CC protein-coupled receptor protein. This peptide is used as an antigen to  
CC prepare rabbit anti-bovine 19p2-L31 antibodies which are used in binding  
CC assays. Pharmaceutical compositions containing this ligand may be used  
CC as a pituitary function modulator, a central nervous system modulator  
CC or a pancreatic function modulator. This ligand could have specific  
CC applications as a prophylactic or therapeutic agent for dementia,  
CC depression, hyperkinetic syndrome, disturbance of consciousness, anxiety  
CC syndrome, schizophrenia, trauma, growth hormone secretory disease,  
CC hyper- and polypthagia, hyperlipidaemia, hypercholesterolaemia,  
CC hyperglyceridaemia, hyperprolactinaemia, diabetes, cancer, pancreatitis,  
CC renal disease, Turner's syndrome, neurosis, asthma, rheumatoid arthritis,  
CC spinal injury, transient brain ischaemia, epilepsy, amyotrophic lateral  
CC sclerosis, acute myocardial infarction, infertility, spinocerebellar  
CC degeneration, bone fracture, trauma, atopic dermatitis, osteoporosis  
CC and/or oligogalacta. Assays can also be developed to screen compounds  
CC which are capable of altering the binding activity of the ligand  
CC affecting activation of the G protein-coupled receptor protein.  
XX  
XX Sequence 15 AA:  
SQ

Query Match 97.1%; Score 33; DB 18; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 REVGRF 6  
| | | | |  
DB 10 rpvgrf 15

RESULT 4  
AAW97230  
ID AAW97230 standard; Peptide: 15 AA.  
XX  
XX AAW97230;  
AC  
XX  
XX 06-MAY-1999 (first entry)  
DT  
XX  
XX C-terminal ligand polypeptide derived antigen.  
DE  
XX  
XX G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation; modulation;  
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dysospermia;  
KW contraceptive; placental function; choriocarcinoma; hydatid mole;  
KW Irruption mole; abortion; unfertility fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia; prolactin secretion.  
XX  
XX Synthetic.  
OS  
XX  
XX W09858962-A1.  
PN  
XX  
XX 30-DEC-1998.  
PD  
XX  
XX 22-JUN-1998; 98WO-JP02765.  
PF  
XX  
XX 23-JUN-1997; 97JP-0165437.  
PR

XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA  
XX  
XX Fuji R, Hinuma S, Kawamata Y, Matsumoto H;  
PI  
XX  
XX WPI: 1999-105614/09.  
DR  
XX  
XX Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PI syndrome, tumours, autoimmune disease or abnormal pregnancy  
XX  
XX  
XX Example 43; Page 123; 241pp; English.  
XX  
XX AAW97229-31 represent a ligand polypeptide derived fragments used to  
CC produce antibodies. The specification describes an agent for modulating  
CC prolactin secretion which comprises a ligand polypeptide or a salt, for  
CC a G protein-coupled receptor (GPCR) protein. The agents for promoting  
CC prolactin secretion can be used for treating or preventing  
CC hypovarianism, gonocyst cacogenesis, menopausal syndrome, euthyroid or  
CC hypometabolism. They can be used for promoting lactation in a domestic  
CC mammal and as an aphrodisiac. The agents for inhibiting prolactin  
CC secretion can be used for treating or preventing pituitary adenomatosis,  
CC brain tumour, emmenopathy, autoimmune disease, prolactinoma,  
CC infertility, impotence, amenorrhea, galactorrhea, acromegaly,  
CC Chiari-Frommel syndrome, Argonz-del Castillo syndrome, Forbes-Albright  
CC syndrome, lymphoma, Sheehan syndrome or dysospermia. The inhibitory  
CC agents can also be used as contraceptives. The agents for modulating  
CC placental function can be used for treating or preventing  
CC choriocarcinoma, hydatid mole, irruption mole, abortion, unfertility  
CC fetus, abnormal saccharometabolism, abnormal lipidmetabolism or  
CC oxytocia.  
XX  
XX Sequence 15 AA:  
SQ

Query Match 97.1%; Score 33; DB 20; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 REVGRF 6  
| | | | |  
DB 10 rpvgrf 15

RESULT 5  
AAV49293  
ID AAV49293 standard; peptide: 15 AA.  
XX  
XX AAV49293;  
AC  
XX  
XX 22-FEB-2000 (first entry)  
DT  
XX  
XX 19p2 ligand peptide fragment.  
DE  
XX  
XX Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
KW  
XX  
XX Homo sapiens.  
OS  
XX  
XX Key Location/Qualifiers  
FH Modified-site 15 /note="C-terminal amide"  
FT  
XX  
XX W09960112-A1.  
PN  
XX  
XX 25-NOV-1999.  
PD  
XX  
XX 20-MAY-1999; 99WO-JP02650.  
PF  
XX  
XX 21-MAY-1998; 98JP-0140293.  
PR  
XX  
XX (TAKE ) TAKEDA CHEM IND LTD.  
PA

PI Matsumoto H, Kitada C, Hinuma S;  
XX WPI; 2000-039381/03.  
XX  
XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
XX disclosure; Page 26; 73pp; Japanese.  
XX  
XX The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA;

Query Match 97.1%; Score 33; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||  
Db 10 rpvgrf 15

RESULT 6  
AAY49295 standard; peptide; 15 AA.  
XX  
AC AAY49295;  
XX

DT 22-FEB-2000 (first entry)  
XX

DE 19p2 ligand peptide fragment.  
XX

KM Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX

OS Homo sapiens.  
XX

Key Location/Qualifiers  
FH Modified-site 1  
FT Modified-site 4 /note- "N-terminal acetylation"  
FT Modified-site 15 /note- "acetylated Tyr".  
FT Modified-site 15 /note- "C-terminal amide"

WO9960112-A1.  
XX

PD 25-NOV-1999.  
XX

PF 20-MAY-1999; 99MO-JP02650.  
XX

PR 21-MAY-1998; 98JP-0140293.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Matsumoto H, Kitada C, Hinuma S;  
XX

DR WPI; 2000-039381/03.  
XX

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
XX disclosure; Page 26; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA;

Query Match 97.1%; Score 33; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||  
Db 10 rpvgrf 15

RESULT 7  
AAY49296 standard; peptide; 15 AA.  
XX  
AC AAY49296;  
XX

DT 22-FEB-2000 (first entry)  
XX

DE 19p2 ligand peptide fragment.  
XX

KM Monoclonal antibody; 19p2 ligand; diagnosis; prolactin secretion;  
KW pituitary; regulatory mechanism; central nervous system; pancreatic.  
XX

OS Homo sapiens.  
XX

Key Location/Qualifiers  
FH Modified-site 15  
FT Modified-site 15 /note- "C-terminal amide"

WO9960112-A1.  
XX

PD 25-NOV-1999.  
XX

PF 20-MAY-1999; 99MO-JP02650.  
XX

PR 21-MAY-1998; 98JP-0140293.  
XX

PA (TAKE ) TAKEDA CHEM IND LTD.  
XX

PI Matsumoto H, Kitada C, Hinuma S;  
XX

DR WPI; 2000-039381/03.  
XX

XX New monoclonal antibodies, useful in diagnosis, as drugs and in  
PT studying diseases related to ligand abnormality -  
XX  
XX disclosure; Page 27; 73pp; Japanese.

CC The invention provides a monoclonal antibody which has a specific  
CC reaction with the part peptide of the C-terminal of 19p2 ligand or its  
CC derivative. The antibodies can be used in diagnosis or to treat or  
CC prevent diseases associated with abnormality in the pituitary function  
CC regulatory mechanism (e.g. promotion of prolactin secretion), central  
CC nervous regulatory mechanism, and pancreatic function regulatory  
CC mechanism. The antibody-based immunoassay can also be applied in  
CC clarifying the physiological functions of the ligand and its derivative.  
CC Sequences AAY49290-302 represent peptide fragments of the 19p2 ligand.  
XX  
SQ Sequence 15 AA;

Query Match 97.1%; Score 33; DB 21; Length 15;  
Best Local Similarity 100.0%; Pred. No. 1.4;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RPVGRF 6  
Db 10 rpvgrf 15

RESULT 8  
AAW31394  
ID AAW31394 standard; Peptide: 20 AA.  
XX  
AC AAW31394;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Human type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Homo sapiens.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB; AAV02431.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 185; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel human type  
CC ligand polypeptide corresponding to amino acid residues 34 to 53 of the  
CC sequence represented in AAW31390 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 97.1%; Score 33; DB 18; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RPVGRF 6  
Db 15 rpvgrf 20

RESULT 9  
AAW31387  
ID AAW31387 standard; Peptide: 20 AA.  
XX  
AC AAW31387;  
XX  
DT 06-APR-1998 (first entry)  
XX  
DE Rat type G protein-coupled receptor ligand fragment 4.  
XX  
KW G protein-coupled receptor; ligand binding; pharmaceutical;  
KW modulator; pituitary; central nervous system; pancreas; prophylactic;  
KW therapeutic agent.  
XX  
OS Rat sp.  
XX  
PN W09724436-A2.  
XX  
PD 10-JUL-1997.  
XX  
PE 26-DEC-1996; 96WO-JP03821.  
XX  
PR 18-SEP-1996; 96JP-0246573.  
PR 28-DEC-1995; 95JP-0343371.  
PR 15-MAR-1996; 96JP-0059419.  
PR 12-AUG-1996; 96JP-0211805.  
XX  
PA (TAKE ) TAKEDA CHEM IND LTD.  
XX  
PI Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;  
PI Kawamata Y, Kitada C;  
XX  
DR WPI: 1997-363672/33.  
DR N-PSDB; AAV02424.  
XX  
PT Ligand peptide for G protein-coupled receptor - acts by modulating  
PT function in the central nervous system, pancreas and pituitary gland  
XX  
PS Claim 2; Page 180; 258pp; English.  
XX  
CC This sequence represents a peptide fragment from a novel rat type  
CC ligand polypeptide corresponding to amino acid residues 33 to 52 of the  
CC sequence represented in AAW31383 and is used in an assay to monitor  
CC ligand binding to the G protein-coupled receptor protein. Pharmaceutical  
CC compositions containing this ligand may be used as a pituitary function  
CC modulator, a central nervous system modulator or a pancreatic function  
CC modulator. This ligand could have specific applications as a  
CC prophylactic or therapeutic agent for dementia, depression, hyperkinetic  
CC syndrome, disturbance of consciousness, anxiety syndrome, schizophrenia,  
CC trauma, growth hormone secretory disease, hyper- and polyphagia,  
CC hyperlipidaemia, hypercholesterolaemia, hyperglycaemia,  
CC hyperprolactaemia, diabetes, cancer, pancreatitis, renal disease,  
CC Turner's syndrome, neurosis, asthma, rheumatoid arthritis, spinal injury,  
CC transient brain ischaemia, epilepsy, amyotrophic lateral sclerosis,  
CC acute myocardial infarction, infertility, spinocerebellar degeneration,  
CC bone fracture, trauma, atopic dermatitis, osteoporosis and/or  
CC oligosaccharia. Assays can also be developed to screen compounds which are  
CC capable of altering the binding activity of the ligand affecting  
CC activation of the G protein-coupled receptor protein.  
XX  
SQ Sequence 20 AA;

Query Match 97.1%; Score 33; DB 18; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.8;

Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 RPYGRF 6  
 |||||  
 Db 15 rpygrf 20

## RESULT 10

AAW31374  
 ID AAW31374 standard; Peptide: 20 AA.

XX AAW31374;

DT 06-APR-1998 (first entry)

DE Bovine G protein-coupled receptor ligand peptide fragment 4.

KM G protein-coupled receptor; ligand binding; pharmaceutical;

KM modulator; pituitary; central nervous system; pancreas; prophylactic;

OS therapeutic agent.

XX Bos taurus.

PN W09724436-A2.

XX 10-JUL-1997.

XX 26-DEC-1996; 96MO-JP03821.

XX 18-SEP-1996; 96JP-0246573.

PR 28-DEC-1995; 95JP-0343371.

PR 15-MAR-1996; 96JP-0059419.

PR 12-AUG-1996; 96JP-0211805.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Fukusumi S, Habata Y, Hinuma S, Hosoya M;

PI Kawamata Y, Kitada C;

XX WPI; 1997-363672/33.

DR N-PSDB; AAV02397.

XX Ligand peptide for G protein-coupled receptor - acts by modulating

PT function in the central nervous system, pancreas and pituitary gland

XX Claim 2; Page 161; 258pp; English.

XX This sequence represents a peptide fragment of a novel bovine pituitary  
 CC derived ligand corresponding to amino acid residues 34 to 53 of the  
 CC sequence in AAW31368 and is used in an assay to monitor ligand binding  
 CC to the G protein-coupled receptor protein. Pharmaceutical compositions  
 CC containing this ligand may be used as a pituitary function modulator, a  
 CC central nervous system modulator or a pancreatic function modulator.  
 CC This ligand could have specific applications as a prophylactic or  
 CC therapeutic agent for dementia, depression, hyperkinetic syndrome,  
 CC disturbance of consciousness, anxiety syndrome, schizophrenia, trauma,  
 CC growth hormone secretory disease, hyper- and polyphagia,  
 CC hypercholesterolemia, hyperglycemia, hyperlipidemia,  
 CC hyperprolactinemia, diabetes, cancer, pancreatitis, renal disease,  
 CC Turner's syndrome, neurosis, rheumatoid arthritis, spinal injury,  
 CC transient brain ischaemia, amyotrophic lateral sclerosis, acute  
 CC myocardial infarction, spinocerebellar degeneration, bone fracture,  
 CC trauma, atopic dermatitis, osteoporosis, asthma, epilepsy, infertility  
 CC and/or oligosaccharia. Assays can also be developed to screen compounds  
 CC which are capable of altering the binding activity of the ligand thus  
 CC affecting activation of the G protein-coupled receptor protein.

XX Sequence 20 AA;

Query Match 97.1%; Score 33; DB 18; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPYGRF 6  
 |||||  
 Db 15 rpygrf 20

## RESULT 11

AAW97232  
 ID AAW97232 standard; peptide: 20 AA.

XX AAW97232;

DT 06-MAY-1999 (first entry)

XX Bovine pituitary-derived ligand polypeptide fragment.

KM Bovine pituitary-derived ligand; modulation; prolactin secretion;

KM G protein-coupled receptor; GPCR; hypovarianism; gonocyte caecogenesis;

KM menopausal syndrome; euthyroid; hypometabolism; lactation;

KM pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;

KM prolactinoma; infertility; impotence; amenorrhoea; galactorrhea;

KM acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;

KM Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;

KM contraceptive; placental function; choriocarcinoma; hydatid mole;

KM interruption mole; abortion; unfertilized fetus; abnormal saccharometabolism;

KM abnormal lipidmetabolism; oxytocia.

XX Bos sp.

PN W09858962-A1.

XX 30-DEC-1998.

XX 22-JUN-1998; 98MO-JP02765.

XX 23-JUN-1997; 97JP-0165437.

XX (TAKE ) TAKEDA CHEM IND LTD.

XX Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

PI WPI; 1999-105614/09.

XX Use of G protein-coupled receptor ligands - for modulating prolactin

PT secretion or placental function, e.g. for treating menopausal

PT syndrome, tumours, autoimmune disease or abnormal pregnancy

XX Claim 3; Page 136; 241pp; English.

XX The present sequence represents a bovine pituitary-derived ligand  
 CC fragment. It is used in the course of the invention. The specification  
 CC describes an agent for modulating prolactin secretion which comprises a  
 CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
 CC protein. The agents for promoting prolactin secretion can be used for  
 CC treating or preventing hypovarianism, gonocyte caecogenesis, menopausal  
 CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
 CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
 CC inhibiting prolactin secretion can be used for treating or preventing  
 CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
 CC prolactinoma, infertility, impotence, amenorrhoea, galactorrhea,  
 CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
 CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
 CC The inhibitory agents can also be used as contraceptives. The agents for  
 CC modulating placental function can be used for treating or preventing  
 CC choriocarcinoma, hydatid mole, interruption mole, abortion, unfertilized fetus,  
 CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX Sequence 20 AA;

Query Match 97.1%; Score 33; DB 20; Length 20;  
 Best Local Similarity 100.0%; Pred. No. 1.8;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
Db 15 rpvgvf 20

## RESULT 12

AAW97234  
ID AAW97234 standard; peptide: 20 AA.

XX  
AC AAW97234;

XX  
DT 06-MAY-1999 (first entry)

XX  
DE Rat type ligand polypeptide fragment.

XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.

XX  
OS Rattus sp.

XX  
PN W09858962-A1.

XX  
PD 30-DEC-1998.

XX  
PE 22-JUN-1998; 98WO-JP02765.

XX  
PR 23-JUN-1997; 97JP-0165437.

XX  
PA (TAKE ) TAKEDA CHEM IND LTD.

XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX  
DR WPI: 1999-105614/09.

XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy

XX  
PS Claim 3; Page 154; 24pp; English.

XX  
CC The present sequence represents a rat type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydralid mole, interruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX  
SQ Sequence 20 AA;

Query Match 97.1%; Score 33; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
Db 15 rpvgvf 20

## RESULT 13

AAW97236  
ID AAW97236 standard; peptide: 20 AA.

XX  
AC AAW97236;

XX  
DT 06-MAY-1999 (first entry)

XX  
DE Human type ligand polypeptide fragment.

XX  
KW Rat type ligand; modulation; prolactin secretion;  
KW G protein-coupled receptor; GPCR; hypovarianism; gonocyst cacogenesis;  
KW menopausal syndrome; euthyroid; hypometabolism; lactation;  
KW pituitary adenomatosis; brain tumour; emmenopathy; autoimmune disease;  
KW prolactinoma; infertility; impotence; amenorrhea; galactorrhea;  
KW acromegaly; Chiari-Frommel syndrome; Argonz-del Castillo syndrome;  
KW Forbes-Albright syndrome; lymphoma; Sheehan syndrome; dyszoospermia;  
KW contraceptive; placental function; choriocarcinoma; hydralid mole;  
KW interruption mole; abortion; unthrifty fetus; abnormal saccharometabolism;  
KW abnormal lipidmetabolism; oxytocia.

XX  
OS Homo sapiens.

XX  
PN W09858962-A1.

XX  
PD 30-DEC-1998.

XX  
PE 22-JUN-1998; 98WO-JP02765.

XX  
PR 23-JUN-1997; 97JP-0165437.

XX  
PA (TAKE ) TAKEDA CHEM IND LTD.

XX  
PI Fujii R, Hinuma S, Kawamata Y, Matsumoto H;

XX  
DR WPI: 1999-105614/09.

XX  
PT Use of G protein-coupled receptor ligands - for modulating prolactin  
PT secretion or placental function, e.g. for treating menopausal  
PT syndrome, tumours, autoimmune disease or abnormal pregnancy

XX  
PS Claim 3; Page 166; 24pp; English.

XX  
CC The present sequence represents a human type ligand fragment. It  
CC is used in the course of the invention. The specification describes  
CC an agent for modulating prolactin secretion which comprises a  
CC ligand polypeptide or a salt, for a G protein-coupled receptor (GPCR)  
CC protein. The agents for promoting prolactin secretion can be used for  
CC treating or preventing hypovarianism, gonocyst cacogenesis, menopausal  
CC syndrome, euthyroid or hypometabolism. They can be used for promoting  
CC lactation in a domestic mammal and as an aphrodisiac. The agents for  
CC inhibiting prolactin secretion can be used for treating or preventing  
CC pituitary adenomatosis, brain tumour, emmenopathy, autoimmune disease,  
CC prolactinoma, infertility, impotence, amenorrhea, galactorrhea,  
CC acromegaly, Chiari-Frommel syndrome, Argonz-del Castillo syndrome,  
CC Forbes-Albright syndrome, lymphoma, Sheehan syndrome or dyszoospermia.  
CC The inhibitory agents can also be used as contraceptives. The agents for  
CC modulating placental function can be used for treating or preventing  
CC choriocarcinoma, hydralid mole, interruption mole, abortion, unthrifty fetus,  
CC abnormal saccharometabolism, abnormal lipidmetabolism or oxytocia.

XX  
SQ Sequence 20 AA;

Query Match 97.1%; Score 33; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
111111  
Db 15 rpvgrf 20

## RESULT 14

AAW95191  
ID AAW95191 standard; peptide; 20 AA.

AC AAW95191;

DT 10-MAR-1999 (first entry)

DE Bovine pituitary-derived ligand polypeptide fragment.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
XX GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; bovine.

OS Bos sp.

PN WO9849295-A1.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

DR WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

PS Example 19; Page 151; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of the  
CC polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop transgenic  
CC animals. The present sequence represents a bovine genome-derived ligand  
CC polypeptide fragment which is similar to the murine ligand-polypeptide.

XX Sequence 20 AA;

Query Match 97.1%; Score 33; DB 20; Length 20;

Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
111111  
Db 15 rpvgrf 20

## RESULT 15

AAW95175  
ID AAW95175 standard; Protein; 20 AA.

AC AAW95175;

DT 10-MAR-1999 (first entry)

DE Murine pituitary-derived ligand polypeptide antigenic epitope.

XX Pituitary-derived ligand polypeptide; G-protein coupled orphan receptor;  
XX GPR10; UHR-1; modulator; pituitary; central nervous system; pancreas;  
KM tissue; screen; therapeutic; binding; senile dementia; ligand; murine;  
KM Alzheimer's disease; Parkinson's disease; Huntington's disease; drug;  
KM Creutzfeld-Jakob disease; poisoning; schizophrenia; growth hormone;  
KM secretion; diabetes; cancer; rheumatoid arthritis; epilepsy; vasopressor;  
KM gene therapy; transgenic animal; epitope.

OS Mus sp.

PN WO9849295-A1.

PD 05-NOV-1998.

PF 27-APR-1998; 98WO-JP01923.

PR 28-APR-1997; 97JP-0109974.

PA (TAKE ) TAKEDA CHEM IND LTD.

PI Fukusumi S, Hinuma S;

DR WPI; 1999-009423/01.

XX New polypeptide ligand for orphan G protein coupled receptors - used  
PT for treating disorders of central nervous system, pituitary and  
PT pancreas, and for drug screening

PS Disclosure; Page 26; 206pp; English.

XX The invention relates to a murine pituitary-derived ligand polypeptide  
CC which is a ligand for the G-protein coupled orphan receptor designated  
CC GPR10 (human) or UHR-1 (rat). Cells transformed with a vector containing  
CC the ligand polypeptide encoding DNA are used to produce a recombinant  
CC ligand polypeptide. The ligand polypeptide, and its fragments, modulate  
CC function of the pituitary, central nervous system, pancreas and other  
CC tissues and can be used to screen for agents that modulate binding of  
CC the polypeptide to the receptor; to quantify the amount of receptor in a  
CC sample and to raise antibodies. They may also be used therapeutically,  
CC e.g. to treat senile dementia; Alzheimer's, Parkinson's or Huntington's  
CC diseases; Creutzfeld-Jakob disease; poisoning by heavy metals or drugs;  
CC diabetes; schizophrenia; disorders of growth hormone secretion; cancer;  
CC rheumatoid arthritis; epilepsy and many others, also to improve post-  
CC operative nutritional status and as vasopressor. Transgenic animals  
CC carrying the ligand polypeptide encoding DNA or its mutin are used to  
CC study the function of the polypeptide-expressing genes, as models of  
CC disease, for drug screening and as source of cell lines. The ligand  
CC polypeptide DNA is used as a source of probes and primers; to identify  
CC related sequences; in receptor-binding assays; for production of Ab and  
CC antisera; in drug development; for gene therapy and to develop  
CC transgenic animals. Sequences AAW95174 to AAW95178 represent antigenic  
CC epitopes which can be used for the preparation of anti-ligand polypeptide  
CC antibody.

XX Sequence 20 AA;



Query Match 97.1%; Score 33; DB 20; Length 20;  
Best Local Similarity 100.0%; Pred. No. 1.8;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPVGRF 6  
|||||  
Db 15 rpvgrf 20

Search completed: September 13, 2002, 09:18:36  
Job time: 501 sec

---

THIS PAGE BLANK (USPTO)

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using SW model

Run on: September 13, 2002, 09:20:59 ; Search time 136.62 Seconds  
(without alignments)  
1.251 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_15\_21

Perfect score: 34

Sequence: 1 RPYGRFX 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 231628 seqs, 24425594 residues

Total number of hits satisfying chosen parameters: 231628

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Issued\_Patents\_AA.\*

1: /cgn2\_6/ptodata/2/1aa/5A.COMB.pep:\*  
2: /cgn2\_6/ptodata/2/1aa/5B.COMB.pep:\*  
3: /cgn2\_6/ptodata/2/1aa/6A.COMB.pep:\*  
4: /cgn2\_6/ptodata/2/1aa/6B.COMB.pep:\*  
5: /cgn2\_6/ptodata/2/1aa/PCTUS.COMB.pep:\*  
6: /cgn2\_6/ptodata/2/1aa/Backfilest1.pep:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	97.1	15	4	US-08-776-971-93 Sequence 93, Appl
2	33	97.1	20	3	US-09-105-678A-34 Sequence 34, Appl
3	33	97.1	20	3	US-09-105-678A-40 Sequence 40, Appl
4	33	97.1	20	3	US-09-105-678A-46 Sequence 46, Appl
5	33	97.1	20	4	US-08-776-971-8 Sequence 8, Appl
6	33	97.1	20	4	US-08-776-971-50 Sequence 50, Appl
7	33	97.1	20	4	US-08-776-971-64 Sequence 64, Appl
8	33	97.1	20	4	US-08-776-971-98 Sequence 98, Appl
9	33	97.1	20	4	US-09-421-208-34 Sequence 34, Appl
10	33	97.1	20	4	US-09-421-208-40 Sequence 40, Appl
11	33	97.1	20	4	US-09-421-208-46 Sequence 46, Appl
12	33	97.1	21	3	US-09-105-678A-28 Sequence 28, Appl
13	33	97.1	21	3	US-09-105-678A-35 Sequence 35, Appl
14	33	97.1	21	3	US-09-105-678A-41 Sequence 41, Appl
15	33	97.1	21	3	US-09-105-678A-47 Sequence 47, Appl
16	33	97.1	21	4	US-08-776-971-9 Sequence 9, Appl
17	33	97.1	21	4	US-08-776-971-51 Sequence 51, Appl
18	33	97.1	21	4	US-08-776-971-65 Sequence 65, Appl
19	33	97.1	21	4	US-09-421-208-28 Sequence 28, Appl
20	33	97.1	21	4	US-09-421-208-35 Sequence 35, Appl
21	33	97.1	21	4	US-09-421-208-41 Sequence 41, Appl
22	33	97.1	21	4	US-09-421-208-47 Sequence 47, Appl
23	33	97.1	22	3	US-09-105-678A-36 Sequence 36, Appl
24	33	97.1	22	3	US-09-105-678A-42 Sequence 42, Appl
25	33	97.1	22	3	US-09-105-678A-48 Sequence 48, Appl
26	33	97.1	22	4	US-08-776-971-10 Sequence 10, Appl
27	33	97.1	22	4	US-08-776-971-52 Sequence 52, Appl

28	33	97.1	22	4	US-08-776-971-66 Sequence 66, Appl
29	33	97.1	22	4	US-08-776-971-73 Sequence 73, Appl
30	33	97.1	22	4	US-09-421-208-36 Sequence 36, Appl
31	33	97.1	22	4	US-09-421-208-42 Sequence 42, Appl
32	33	97.1	22	4	US-09-421-208-48 Sequence 48, Appl
33	33	97.1	31	3	US-09-105-678A-7 Sequence 7, Appl
34	33	97.1	31	3	US-09-105-678A-8 Sequence 8, Appl
35	33	97.1	31	3	US-09-105-678A-9 Sequence 9, Appl
36	33	97.1	31	3	US-09-105-678A-31 Sequence 31, Appl
37	33	97.1	31	3	US-09-105-678A-37 Sequence 37, Appl
38	33	97.1	31	3	US-09-105-678A-43 Sequence 43, Appl
39	33	97.1	31	4	US-09-172-353-4 Sequence 4, Appl
40	33	97.1	31	4	US-08-776-971-5 Sequence 5, Appl
41	33	97.1	31	4	US-08-776-971-47 Sequence 47, Appl
42	33	97.1	31	4	US-08-776-971-61 Sequence 61, Appl
43	33	97.1	31	4	US-08-776-971-97 Sequence 97, Appl
44	33	97.1	31	4	US-09-421-208-7 Sequence 7, Appl
45	33	97.1	31	4	US-09-421-208-8 Sequence 8, Appl

#### ALIGNMENTS

RESULT 1  
US-08-776-971-93  
Sequence 93, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPEPTIDES, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776, 971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 93:  
SEQUENCE CHARACTERISTICS:

LENGTH: 15 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 93:  
US-08-776-971-93

Query Match 97.1% Score 33; DB 4; Length 15;  
Best Local Similarity 100.0%; Pred. No. 0.39;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
Db 10 RPVGRF 15

RESULT 2  
US-09-105-678A-34  
Sequence 34, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-34

Query Match 97.1% Score 33; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
Db 15 RPVGRF 20

RESULT 3  
US-09-105-678A-40  
Sequence 40, Application US/09105678A  
Patent No. 6103882

GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

INFORMATION FOR SEQ ID NO: 40:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-40

Query Match 97.1% Score 33; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
Db 15 RPVGRF 20

RESULT 4  
US-09-105-678A-46  
Sequence 46, Application US/09105678A  
Patent No. 6103882

GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440

ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-46

Query Match 97.1%; Score 33; DB 3; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
DB 15 RPVGRF 20

RESULT 5  
US-08-776-971-8  
Sequence 8, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995

APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 8:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 8:  
US-08-776-971-8

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||||  
DB 15 RPVGRF 20

RESULT 6  
US-08-776-971-50  
Sequence 50, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fujii, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996

APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 50:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 50:  
US-08-776-971-50

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
Db 15 RPYGRF 20

RESULT 7  
US-08-776-971-64  
Sequence 64, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.

REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 64:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
FRAGMENT TYPE: internal  
SEQUENCE DESCRIPTION: SEQ ID NO: 64:  
US-08-776-971-64

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
Db 15 RPYGRF 20

RESULT 8  
US-08-776-971-98  
Sequence 98, Application US/08776971B  
Patent No. 6228984  
GENERAL INFORMATION:  
APPLICANT: Hinuma, Shuji  
Habata, Yugo  
Kawamata, Yuji  
Hosoya, Masaki  
Fuji, Ryo  
Fukusumi, Shoji  
Kitada, Chieko  
TITLE OF INVENTION: POLYPROTEINS, THEIR PRODUCTION AND USE  
NUMBER OF SEQUENCES: 140  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FastSeq for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/776,971B  
FILING DATE: 06-Feb-1997  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: PCT/JP96/03821  
FILING DATE: 28-DEC-1996  
APPLICATION NUMBER: JP 7/343371  
FILING DATE: 28-DEC-1995  
APPLICATION NUMBER: JP 8/59419  
FILING DATE: 15-MAR-1996  
APPLICATION NUMBER: JP 8/211805  
FILING DATE: 12-AUG-1996  
APPLICATION NUMBER: JP 8/246573  
FILING DATE: 18-SEP-1996  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 47176  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400

TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 98:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; SEQUENCE DESCRIPTION: SEQ ID NO: 98:  
US-08-776-971-98

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|11111  
DB 15 RPVGRF 20

RESULT 9  
US-09-421-208-34  
; Sequence 34, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 34:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-34

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
OY 1 RPVGRF 6

DB 15 RPVGRF 20  
|11111

RESULT 10  
US-09-421-208-40  
; Sequence 40, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
; STREET: 130 Water Street  
; CITY: Boston  
; STATE: MA  
; COUNTRY: USA  
; ZIP: 02109  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/421,208  
; FILING DATE:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 09/105,678  
; FILING DATE: 26-JUN-1998  
; APPLICATION NUMBER: JP 172118/1997  
; FILING DATE: 27-JUN-1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Conlin, David G.  
; REGISTRATION NUMBER: 27,026  
; REFERENCE/DOCKET NUMBER: 48466-342  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 617-523-3400  
; TELEFAX: 617-523-6440  
; INFORMATION FOR SEQ ID NO: 40:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 20 amino acids  
; TYPE: amino acid  
; STRANDEDNESS:  
; TOPOLOGY: linear  
; MOLECULE TYPE: peptide  
US-09-421-208-40

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|11111  
DB 15 RPVGRF 20

RESULT 11  
US-09-421-208-46  
; Sequence 46, Application US/09421208  
; Patent No. 6258561  
; GENERAL INFORMATION:  
; APPLICANT: Suenaga, Masato  
; APPLICANT: Moriya, Takeo  
; APPLICANT: Tanaka, Yoko  
; APPLICANT: Nishimura, Osamu  
; TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
US-09-421-208-46

ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/421,208  
FILING DATE:  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 09/105,678  
FILING DATE: 26-JUN-1998  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 46:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 20 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-421-208-46

Query Match 97.1%; Score 33; DB 4; Length 20;  
Best Local Similarity 100.0%; Pred. No. 0.52;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
Db 15 RPVGRF 20

RESULT 12  
US-09-105-678A-28  
Sequence 28, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997

FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 28:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 10  
OTHER INFORMATION: /product- "Ala or Thr"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 11  
OTHER INFORMATION: /product- "Gly or Ser"  
FEATURE:  
NAME/KEY: Modified-site  
LOCATION: 21  
OTHER INFORMATION: /product- "Gly-Oh or Gly-Arg"  
US-09-105-678A-28

Query Match 97.1%; Score 33; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
Db 15 RPVGRF 20

RESULT 13  
US-09-105-678A-35  
Sequence 35, Application US/09105678A  
Patent No. 6103882  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105,678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:



TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 35:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-35

Query Match 97.1%; Score 33; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
DB 15 RPVGRF 20

RESULT 14  
US-09-105-678A-41  
Sequence 41, Application US/09105678A  
Patent No. 6103862  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105.678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 41:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-41

Query Match 97.1%; Score 33; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
DB 15 RPVGRF 20

DB 15 RPVGRF 20

RESULT 15  
US-09-105-678A-47  
Sequence 47, Application US/09105678A  
Patent No. 6103862  
GENERAL INFORMATION:  
APPLICANT: Suenaga, Masato  
APPLICANT: Moriya, Takeo  
APPLICANT: Tanaka, Yoko  
APPLICANT: Nishimura, Osamu  
TITLE OF INVENTION: METHOD OF PRODUCING A 19P2 LIGAND  
NUMBER OF SEQUENCES: 52  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: DIKE, BRONSTEIN, ROBERTS & CUSHMAN, LLP  
STREET: 130 Water Street  
CITY: Boston  
STATE: MA  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/09/105.678A  
FILING DATE: 26-JUN-1998  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: JP 172118/1997  
FILING DATE: 27-JUN-1997  
ATTORNEY/AGENT INFORMATION:  
NAME: Conlin, David G.  
REGISTRATION NUMBER: 27,026  
REFERENCE/DOCKET NUMBER: 48466-342  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-523-3400  
TELEFAX: 617-523-6440  
INFORMATION FOR SEQ ID NO: 47:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 21 amino acids  
TYPE: amino acid  
STRANDEDNESS:  
TOPOLOGY: linear  
MOLECULE TYPE: peptide  
US-09-105-678A-47

Query Match 97.1%; Score 33; DB 3; Length 21;  
Best Local Similarity 100.0%; Pred. No. 0.55;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
DB 15 RPVGRF 20

Search completed: September 13, 2002, 09:20:59  
Job time: 624 sec

**THIS PAGE BLANK (USPTO)**

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:24:04 ; Search time 172.41 seconds

(without alignments)  
3.901 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_15\_21

Perfect score: 34

Sequence: 1 RPYGRFX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283138 seqs, 96089334 residues

Total number of hits satisfying chosen parameters: 283138

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

1: p1r1:\*  
2: p1r2:\*  
3: p1r3:\*  
4: p1r4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	97.1	83	2	JC1607
2	33	97.1	445	2	T35893
3	33	97.1	511	2	T40334
4	30	88.2	107	2	AC3538
5	30	88.2	149	2	AG2262
6	30	88.2	154	2	T34825
7	30	88.2	173	2	S76779
8	30	88.2	501	2	S66763
9	30	88.2	559	2	G83897
10	29	85.3	170	2	S76067
11	29	85.3	172	2	T02229
12	29	85.3	226	2	A65057
13	29	85.3	289	1	A37209
14	29	85.3	308	2	AC1053
15	29	85.3	315	2	T46156
16	29	85.3	362	2	G96735
17	29	85.3	367	2	T04521
18	29	85.3	379	2	T45286
19	28	82.4	138	2	S24102
20	28	82.4	193	2	T44148
21	28	82.4	239	2	AB1887
22	28	82.4	246	2	H98292
23	28	82.4	246	2	T04521
24	28	82.4	251	2	AB3390
25	28	82.4	257	2	F83104
26	28	82.4	260	2	T36846
27	28	82.4	294	2	F84922
28	28	82.4	299	2	C70643
29	28	82.4	299	2	B86770

30	28	82.4	300	1	S08244	conserved hypothet
31	28	82.4	305	2	G96962	UDP-N-acetylennolpy
32	28	82.4	316	2	G95161	UDP-N-acetylennolpy
33	28	82.4	316	2	P98027	UDP-N-acetylennolpy
34	28	82.4	325	2	P98292	ribose ABC transpo
35	28	82.4	325	2	AE2991	ABC transporter, m
36	28	82.4	367	2	T44687	cobalamin biosynth
37	28	82.4	409	2	T47470	3-ketoadyl-CoA deh
38	28	82.4	424	2	D69399	hypothetical prote
39	28	82.4	433	2	BA7041	hypothetical prote
40	28	82.4	457	2	AC2728	polysaccharide Dio
41	28	82.4	458	2	AC2935	conserved hypothet
42	28	82.4	497	2	D98347	hypothetical prote
43	28	82.4	522	2	H97509	probable glycosyl
44	28	82.4	580	2	A96683	hypothetical prote
45	28	82.4	585	2	C69336	probable electron

#### ALIGNMENTS

RESULT 1  
JC7607  
prolactin-releasing peptide - rat  
C:Species: Rattus norvegicus (Norway rat)  
C:Date: 30-Jun-2001 #sequence\_revision 30-Jun-2001 #text\_change 30-Jun-2001  
C:Accession: JC7607  
R:Yamada, M.; Ozawa, A.; Ishii, S.; Shibusawa, N.; Hashida, T.; Hosoya, A.; Biochem. Biophys. Res. Commun. 281: 53-56, 2001  
A:Title: Isolation and characterization of the rat prolactin-releasing peptide gene:  
A:Reference number: JC7607; MUID:21092785; PMID:11178959  
A:Contents: Spleen  
A:Accession: JC7607  
A:Molecule type: DNA  
A:Residues: 1-83 <YAM>  
A:Cross-references: DDBJ:AB040612; DDBJ:AB040613  
C:Comment: This peptide induces arachidonic acid metabolite release from rat anterior release, and stimulation of ACTH secretion from the pituitary.  
C:Genetics:  
A:Gene: PRP  
A:Introns: 33/1

Query Match 97.1%, Score 33; DB 2; Length 83;  
Best Local Similarity 100.0%; Pred. No. 3.5;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPYGRFX 6  
DB 47 RPYGRFX 52

RESULT 2  
T35893  
FAD-dependent oxidoreductase - Streptomyces coelicolor  
C:Species: Streptomyces coelicolor  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 21-Jan-2000  
C:Accession: T35893  
R:Olliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, November 1997  
A:Reference number: Z21592  
A:Accession: T35893  
A:Status: preliminary; translated from GB/EMBL/DDBJ  
A:Molecule type: DNA  
A:Residues: 1-445 <OLI>  
A:Cross-references: EMBL:AL009204; PIDN:CAAI5814.1; GSPDB:GN00070; SCOEDB:SC9B10.24c  
C:Experimental source: strain A3(2)  
C:Genetics:  
A:Gene: SCOEDB:SC9B10.24c  
C:Superfamily: L-gulonolactone oxidase

Query Match 97.1%, Score 33; DB 2; Length 445;

Best Local Similarity 100.0%; Pred. No. 18;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPYGRF 6  
|||||  
Db 248 RPYGRF 253

RESULT 3  
T40334  
hypothetical protein SPBC3B8.06 - fission yeast (*Schizosaccharomyces pombe*)  
C:Species: *Schizosaccharomyces pombe*  
C:Date: 03-Dec-1999 #sequence\_revision 03-Dec-1999 #text\_change 03-Dec-1999  
C:Accession: T40334  
R:Lyne, M.; Rajandream, M.A.; Barrell, B.G.; Beck, A.; Reinhardt, R.; Pohl, T.  
submitted to the EMBL Data Library, March 1998  
A:Reference number: Z21921  
A:Accession: T40334  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-511 <LYN>  
A:Cross-references: EMBL:AL022244; PIDN:CAA18295.1; GSPDB:GN00067; SPDB:SPBC3B8.06  
A:Experimental source: strain 972h-; cosmid c3B8  
C:Genetics:  
A:Gene: SPBC3B8.06  
A:Map position: 2

Query Match 97.1%; Score 33; DB 2; Length 511;  
Best Local Similarity 100.0%; Pred. No. 21;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPYGRF 6  
|||||  
Db 95 RPYGRF 100

RESULT 4  
AC3538  
hypothetical protein BMEI10229 [Imported] - *Brucella melitensis* (strain 16M)  
C:Species: *Brucella melitensis*  
C:Date: 01-Feb-2002 #sequence\_revision 01-Feb-2002 #text\_change 01-Feb-2002  
C:Accession: AC3538  
R:Delvecchio, V.G.; Kapetral, V.; Redkar, R.J.; Patra, G.; Mijer, C.; Los, T.; Ivanova,  
Proc. Natl. Acad. Sci. U.S.A. 99, 443-448, 2002  
A:Title: The genome sequence of the facultative intracellular pathogen *Brucella melitensis*  
A:Reference number: AD3252; PMID:11756688  
A:Accession: AC3538  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-107 <KUR>  
A:Cross-references: GB:AE008918; PIDN:AAL53470.1; PID:g17984372; GSPDB:GN00191  
C:Genetics:  
A:Gene: BMEI10229  
A:Map position: 11

Query Match 88.2%; Score 30; DB 2; Length 107;  
Best Local Similarity 83.3%; Pred. No. 20;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPYGRF 6  
|||||  
Db 99 RPYGRF 104

RESULT 5  
AH2262  
hypothetical protein alr3655 [Imported] - *Anabaena* sp. (strain PCC 7120)  
C:Species: *Anabaena* sp.  
A:Note: *Anabaena* sp. (strain PCC 7120) is a synonym of *Nostoc* sp. strain PCC 7120

C:Date: 14-Dec-2001 #sequence\_revision 14-Dec-2001 #text\_change 11-Jan-2002  
C:Accession: AH2262  
R:Kaneo, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Ifugu  
Nakazaki, N.; Shimp, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata  
DNA Res. 8, 205-213, 2001  
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium  
A:Reference number: AB1807; MUID:21595285; PMID:11759840  
A:Accession: AH2262  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-149 <KUR>  
A:Cross-references: GB:BA00019; PIDN:BAW5354.1; PID:g17132788; GSPDB:GN00179  
A:Experimental source: strain PCC 7120  
C:Genetics:  
A:Gene: alr3655

Query Match 88.2%; Score 30; DB 2; Length 149;  
Best Local Similarity 83.3%; Pred. No. 28;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPYGRF 6  
|||||  
Db 7 RPYGRF 12

RESULT 6  
T34825  
hypothetical protein SC2E9.14 SC2E9.14 - *Streptomyces coelicolor*  
C:Species: *Streptomyces coelicolor*  
C:Date: 05-Nov-1999 #sequence\_revision 05-Nov-1999 #text\_change 05-Nov-1999  
C:Accession: T34825  
R:Oliver, K.; Harris, D.; Parkhill, J.; Barrell, B.G.; Rajandream, M.A.  
submitted to the EMBL Data Library, January 1998  
A:Reference number: Z21558  
A:Accession: T34825  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: DNA  
A:Residues: 1-154 <OLI>  
A:Cross-references: EMBL:AL021530; PIDN:CAA16482.1; GSPDB:GN00070; SCOEDB:SC2E9.14  
C:Genetics:  
A:Gene: SCOEDB:SC2E9.14

Query Match 88.2%; Score 30; DB 2; Length 154;  
Best Local Similarity 83.3%; Pred. No. 29;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;  
QY 1 RPYGRF 6  
|||||  
Db 31 RPYGRF 36

RESULT 7  
S76779  
hypothetical protein - *Synechocystis* sp. (strain PCC 6803)  
C:Species: *Synechocystis* sp.  
A:Variety: PCC 6803  
C:Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
C:Accession: S76779  
R:Kaneo, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima,  
O, K.; Okumura, S.; Shimp, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium *Synechocys  
s.*  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76779  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-173 <KAN>  
A:Cross-references: EMBL:090916; GB:AB001339; MUID:g1653715; PIDN:BA18691.1; PID:d101  
A:Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 88.2% Score 30; DB 2; Length 173;  
Best Local Similarity 83.3% Pred. No. 33;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
:|||||  
DB 94 KPVGRF 99

RESULT 8  
S66763  
hypothetical protein YOL070c - yeast (Saccharomyces cerevisiae)  
N:Alternate names: hypothetical protein O1150  
C:Species: Saccharomyces cerevisiae  
C>Date: 12-Jul-1996 #sequence\_revision 12-Jul-1996 #text\_change 05-Nov-1999  
R:Accession: S66763  
R:Alexandarakis, D.; Katsoulou, C.; Tzeremla, M.  
submitted to the Protein Sequence Database, July 1996  
A:Reference number: S66756  
A:Accession: S66763  
A:Molecule type: DNA  
A:Residues: 1-501 <ALE>  
A:Cross-references: EMBL:Z74812, NID:g1419894, PID:e251874, PID:g1419895, GSPDB:GN00015;  
A:Experimental source: strain S288C  
C:Genetics:  
A:Gene: MIPS:YOL070c  
A:Map position: 15L

Query Match 88.2% Score 30; DB 2; Length 501;  
Best Local Similarity 83.3% Pred. No. 92;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
:|||||  
DB 170 KPVGRF 175

RESULT 9  
G83897  
urocanase (urocanate hydratase) hutu [imported] - Bacillus halodurans (strain C-125)  
C:Species: Bacillus halodurans  
C>Date: 01-Dec-2000 #sequence\_revision 01-Dec-2000 #text\_change 15-Jun-2001  
R:Takami, H.; Nakasone, K.; Takaki, Y.; Mieno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hirai  
Nucleic Acids Res. 28, 4317-4331, 2000  
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and  
A:Reference number: A83650; MUID:20512582; PMID:11058132  
A:Accession: G83897  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-559 <STO>  
A:Cross-references: GB:AP001513, GB:BA000004, NID:g10174345, PIDN:BA05702.1; GSPDB:GN00  
A:Experimental source: strain C-125  
C:Genetics:  
A:Gene: hutu  
C:Superfamily: urocanate hydratase

Query Match 88.2% Score 30; DB 2; Length 559;  
Best Local Similarity 83.3% Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
:|||||  
DB 84 KPVGRF 89

RESULT 10  
S76067  
hypothetical protein - Synechocystis sp. (strain PCC 6803)

C:Species: Synechocystis sp.  
A:Variety: PCC 6803  
C>Date: 25-Apr-1997 #sequence\_revision 25-Apr-1997 #text\_change 08-Oct-1999  
R:Accession: S76067  
R:Kaneko, T.; Sato, S.; Kotani, H.; Tanaka, A.; Asamizu, E.; Nakamura, Y.; Miyajima, O. K.; Okumura, S.; Shimpo, S.; Takeuchi, C.; Wada, T.; Watanabe, A.; Yamada, M.; Yas  
DNA Res. 3, 109-136, 1996  
A:Title: Sequence analysis of the genome of the unicellular cyanobacterium Synechocys  
s.  
A:Reference number: S74322; MUID:97061201  
A:Accession: S76067  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-170 <KAN>  
A:Cross-references: EMBL:D63999; GB:AB001339; NID:g1001396; PIDN:BA10045.1; PID:d101  
A>Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1996

Query Match 85.3% Score 29; DB 2; Length 170;  
Best Local Similarity 83.3% Pred. No. 53;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
:|||||  
DB 46 QPVGRF 51

RESULT 11  
T02229  
protein BYJ15 - common tobacco (fragment)  
C:Species: Nicotiana tabacum (common tobacco)  
C>Date: 05-Mar-1999 #sequence\_revision 05-Mar-1999 #text\_change 03-Nov-2000  
R:Accession: T02229  
R:Kojima, H.; Hashizume, K.; Imanishi, S.; Nakamura, K.  
submitted to the EMBL Data Library, July 1997  
A:Description: Early jasmonate-inducible genes of tobacco cells.  
A:Reference number: Z14629  
A:Accession: T02229  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-172 <KOJ>  
A:Cross-references: EMBL:AB005878; NID:d1112079; PID:d1022465  
A:Experimental source: strain Bright yellow 2  
C:Superfamily: myrosinase-associated protein MYAP

Query Match 85.3% Score 29; DB 2; Length 172;  
Best Local Similarity 83.3% Pred. No. 54;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
:|||||  
DB 30 RPTGRF 35

RESULT 12  
A65057  
hypothetical protein b2757 - Escherichia coli (strain K-12)  
C:Species: Escherichia coli  
C>Date: 12-Sep-1997 #sequence\_revision 17-Sep-1997 #text\_change 08-Oct-1999  
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.;  
Science 277, 1453-1462, 1997  
A:Title: The complete genome sequence of Escherichia coli K-12.  
A:Reference number: A64720; MUID:97426617  
A:Accession: A65057  
A:Status: preliminary; nucleic acid sequence not shown; translation not shown  
A:Molecule type: DNA  
A:Residues: 1-226 <BLAT>  
A:Cross-references: GB:AE000359; GB:U00096; NID:g1789110; PIDN:AACT5799.1; PID:g17891  
A:Experimental source: strain K-12, substrain MG1655  
C:Superfamily: Escherichia coli hypothetical protein b2757

Query Match 85.3%; Score 29; DB 2; Length 226;  
Best Local Similarity 83.3%; Pred. No. 70;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
|||  
Db 27 RPVGRF 32

## RESULT 13

A37209  
thiosulfate sulfotransferase (EC 2.8.1.1), hepatic - chicken  
C:Species: Gallus gallus (chicken)  
C:Date: 10-Sep-1999 #sequence\_revision 10-Sep-1999 #text\_change 10-Sep-1999  
C:Accession: A37209  
R:Kohanski, R.A.; Heinrichson, R.L.  
J. Protein Chem. 9, 369-377, 1990  
A:Title: Primary structure of avian hepatic rhodanese.  
A:Reference number: A37209; MVID:91113289  
A:Accession: A37209  
A:Molecule type: protein  
A:Residues: 1-289 <PRO>  
C:Superfamily: thiosulfate sulfotransferase  
C:Keywords: liver; mitochondrion; sulfoprotein; sulfotransferase  
F:244/Active site: Cys (sulfocysteine intermediate) #status predicted

Query Match 85.3%; Score 29; DB 1; Length 289;  
Best Local Similarity 83.3%; Pred. No. 89;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
|||  
Db 182 RPVGRF 187

## RESULT 14

AC1053  
probable membrane protein STY4751 [imported] - Salmonella enterica subsp. enterica serov.  
C:Species: Salmonella enterica subsp. enterica serovar typhi  
A:Note: this species has also been called Salmonella typhi  
C:Date: 09-Nov-2001 #sequence\_revision 09-Nov-2001 #text\_change 09-Nov-2001  
C:Accession: AC1053  
R:Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher, A.;  
Th, T.; Conner, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar, J.;  
Moule, S.; O'Garra, P.  
Nature 413, 848-852, 2001  
A:Authors: Parry, C.; Quail, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.;  
A:Title: Complete genome sequence of a multiple drug resistant Salmonella enterica serov.  
A:Reference number: AB0502; PMID:11677608  
A:Accession: AC1053  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-308 <PAR>  
A:Cross-references: GB:AL513362; PIDN:CAD06872.1; PID:G16505520; GSPDB:GN00176  
C:Genetics:  
A:Gene: STY4751

Query Match 85.3%; Score 29; DB 2; Length 308;  
Best Local Similarity 66.7%; Pred. No. 95;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
|||  
Db 75 RPVGRF 80

## RESULT 15

T46156  
hypothetical protein T4D2.30 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 04-Feb-2000 #sequence\_revision 04-Feb-2000 #text\_change 18-Aug-2000  
C:Accession: T46156  
R:Yakutara, G.; Fartmann, B.; Dauner, D.; Sterr, W.; Holland, R.; Weichselgartner, M.  
submitted to the Protein Sequence Database, December 1999  
A:Reference number: Z23025

A:Accession: T46156  
A>Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-315 <NYA>

A:Cross-references: EMBL:AL132958

A:Experimental source: cultivar Columbia; BAC clone T4D2

C:Genetics:

A:Map position: 3

A:Introns: 78/1; 121/3; 199/3; 285/1

A:Note: T4D2.30

C:Superfamily: myosinase-associated protein MyAP

Query Match 85.3%; Score 29; DB 2; Length 315;  
Best Local Similarity 83.3%; Pred. No. 97;  
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPVGRF 6  
|||  
Db 62 RPVGRF 67

Search completed: September 13, 2002, 09:24:05  
Job time: 780 sec

GenCore version 4.5  
Copyright (c) 1993 - 2000 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:30:49 ; Search time 80.21 Seconds  
(without alignments)  
3.379 Million cell updates/sec

Title: US-09-446-543a-73\_COPY\_15\_21

Perfect score: 34

Sequence: 1 RPYGRFX 7

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 105224 seqs, 38719550 residues

Total number of hits satisfying chosen parameters: 105224

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SwissProt\_40:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	33	97.1	83 1	PRRP_RAT
2	33	97.1	87 1	PRRP_HUMAN
3	33	97.1	98 1	PRRP_BOVIN
4	30	88.2	480 1	GIG1_RHME
5	29	85.3	224 1	YGC1_ECOLI
6	29	85.3	289 1	THTR_CHICK
7	29	85.3	330 1	AR72_HUMAN
8	29	85.3	331 1	AR73_HUMAN
9	28	82.4	292 1	YAD_RHISN
10	28	82.4	299 1	MURB_LACLA
11	28	82.4	300 1	YR33_THEPE
12	28	82.4	526 1	CATB_BRARE
13	28	82.4	847 1	MDOH_ECOLI
14	28	82.4	1057 1	ANPA_MOUSE
15	28	82.4	1057 1	ANPA_RAT
16	28	82.4	1061 1	ANPA_HUMAN
17	28	82.4	1061 1	ANPA_HUMAN
18	27	79.4	11 1	ZAN_PIG
19	27	79.4	66 1	CEP1_ACHFU
20	27	79.4	112 1	IGF2_CHICK
21	27	79.4	140 1	AGN5_APLCA
22	27	79.4	162 1	IPYR_MYCLE
23	27	79.4	162 1	IPYR_MYCLE
24	27	79.4	168 1	YNE2_RHME
25	27	79.4	207 1	YNE8_RHME
26	27	79.4	217 1	RS5_MYCLE
27	27	79.4	233 1	RS4E_HALMA
28	27	79.4	248 1	PSA5_CAEEL
29	27	79.4	250 1	LINC_PSEPA
30	27	79.4	284 1	THR3_MYCTU
31	27	79.4	285 1	CNO7_HUMAN
32	27	79.4	285 1	CNO7_MOUSE
33	27	79.4	292 1	CNT8_HUMAN

34	27	79.4	317 1	EMT_BACHD	Q9A9Y6 bacillus ha
35	27	79.4	332 1	OPT_HUMAN	Q9ABM4 homo sapien
36	27	79.4	338 1	DCUP_AOUAE	Q66667 aquifex aeo
37	27	79.4	355 1	CPD1_DROME	P22058 drosophila
38	27	79.4	394 1	YLP3_PSEPU	P31049 pseudomonas
39	27	79.4	400 1	PCAF_PSEPU	Q51956 pseudomonas
40	27	79.4	413 1	CSD_THEMA	Q9X191 thermotoga
41	27	79.4	491 1	CATB_ORYZA	P29611 oryza sativ
42	27	79.4	492 1	CAT2_CUCPE	P48351 cucurbita p
43	27	79.4	496 1	CAT3_MAIZE	P18123 zea mays (m
44	27	79.4	532 1	CRT1_APHSP	P21134 aphanocapsa
45	27	79.4	552 1	CSF1_MOUSE	P07141 mus musculu

## ALIGNMENTS

RESULT	ID	PRRP_RAT	STANDARD:	PRT:	83 AA.
AC	PRRP_RAT	P81278;			
DT	30-MAY-2000 (Rel. 39, Created)				
DT	30-MAY-2000 (Rel. 39, Last sequence update)				
DT	01-MAR-2002 (Rel. 41, Last annotation update)				
DE	Proactin-releasing peptide precursor (PRRP) (Proactin-releasing hormone) [Contains: Proactin-releasing peptide PrRP31; Proactin-releasing peptide PrRP20].				
DE	releasing peptide PrRP20].				
GN	PRH.				
OS	Rattus norvegicus (Rat).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.				
OX	NCBI_TaxID=10116;				
RN	[1]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Brain;				
RX	MEDLINE=98268781; PubMed=9607765;				
RA	Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,				
RA	Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,				
RA	Kurokawa T., Nishimura O., Onda H., Fujino M.;				
RT	"A prolactin-releasing peptide in the brain.";				
RL	Nature 393:272-276(1998).				
RN	[2]				
RP	TISSUE SPECIFICITY.				
RX	PubMed=10498338;				
RA	Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,				
RA	Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,				
RA	Sumino Y., Fujino M.;				
RT	"Tissue distribution of prolactin-releasing peptide (PRRP) and its				
RT	receptor.";				
RL	Regul. Pept. 83:1-10(1999).				
CC	-I- FUNCTION: Stimulates prolactin (PRL) release and regulates the				
CC	expression of prolactin through its receptor GPR10. May stimulate				
CC	lactotrophs directly to secrete PRL.				
CC	-I- TISSUE SPECIFICITY: Widely expressed, with highest levels in				
CC	medulla oblongata and hypothalamus.				
CC	-----				
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration				
CC	between the Swiss Institute of Bioinformatics and the EMBL outstation -				
CC	the European Bioinformatics Institute. There are no restrictions on its				
CC	use by non-profit institutions as long as its content is in no way				
CC	modified and this statement is not removed. Usage by and for commercial				
CC	entities requires a license agreement (See <a href="http://www.isb-sdb.ch/announce/">http://www.isb-sdb.ch/announce/</a>				
CC	or send an email to <a href="mailto:license@sdb.ch">license@sdb.ch</a> ).				
DR	EMBL: AB015418; BAA29026.1;				
KW	Hormone; Amidation; Signal; Cleavage on pair of basic residues.				
FT	SIGNAL	1	21		
FT	PEPTIDE	22	52		
FT	PEPTIDE	33	52		
FT	MOD. RES	52	52		
SO	SEQUENCE	83 AA;	9215 MW;		

Query Match 97.1%; Score 33; DB 1; Length 83;  
 Best Local Similarity 100.0%; Pred. No. 0.89;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
 DB 47 RPVGRF 52

## RESULT 2

PRRP\_HUMAN STANDARD; PRT; 87 AA.

AC P81277;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP1; Prolactin-releasing peptide PrRP20].  
 DE PRH.

OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.  
 OC NCBI\_TaxID=9606;

NCBI\_TaxID=9606;

SEQUENCE FROM N.A.

TISSUE=Brain;

MDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,

RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,

RA Kurokawa T., Nishimura O., Onda H., Fujino M.;

RT "A prolactin-releasing peptide in the brain.";

RL Nature 393:272-276(1998).

CC [2]

TISSUE SPECIFICITY.

RP PubMed=10498338;

RA Fujii R., Fukusumi S., Hosoya M., Kawamata Y., Habata Y., Hinuma S.,

RA Sekiguchi M., Kitada C., Kurokawa T., Nishimura O., Onda H.,

RA Sumino Y., Fujino M.;

RT "Tissue distribution of prolactin-releasing peptide (PrRP) and its

RT receptor.";

RL Regul. Pept. 83:1-10(1999).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the

CC expression of prolactin through its receptor GPR10. May stimulate

CC lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC

CC

CC

CC

Query Match 97.1%; Score 33; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
 DB 48 RPVGRF 53

## RESULT 3

PRRP\_BOVIN STANDARD; PRT; 98 AA.

AC P81264;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Prolactin-releasing peptide precursor (PrRP) (Prolactin-releasing hormone) [Contains: Prolactin-releasing peptide PrRP1; Prolactin-releasing peptide PrRP20].  
 DE PRH.

OS Bos taurus (Bovine).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;  
 OC Bovidae; Bovinae; Bos.  
 OC NCBI\_TaxID=9913;

NCBI\_TaxID=9913;

SEQUENCE FROM N.A., AND SEQUENCE OF 23-52.

TISSUE=Brain;

MDLINE=98268781; PubMed=9607765;

RA Hinuma S., Habata Y., Fujii R., Kawamata Y., Hosoya M., Fukusumi S.,

RA Kitada C., Masuo Y., Asano T., Matsumoto H., Sekiguchi M.,

RA Kurokawa T., Nishimura O., Onda H., Fujino M.;

RT "A prolactin-releasing peptide in the brain.";

RL Nature 393:272-276(1998).

CC -1- FUNCTION: Stimulates prolactin (PRL) release and regulates the

CC expression of prolactin through its receptor GPR10. May stimulate

CC lactotrophs directly to secrete PRL.

CC -1- TISSUE SPECIFICITY: MEDULLA OBLONGATA AND HYPOTHALAMUS.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

Query Match 97.1%; Score 33; DB 1; Length 98;  
 Best Local Similarity 100.0%; Pred. No. 1.1;  
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
 DB 48 RPVGRF 53

## RESULT 4

GLGI\_RHIME STANDARD; PRT; 480 AA.

AC P58393;  
 DT 01-MAR-2002 (Rel. 41, Created)  
 DT 01-MAR-2002 (Rel. 41, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Glycogen synthase 1 (EC 2.4.1.21) [Starch [bacterial glycogen]

DE synthase 1].

GN GIGAI OR R02846 OR SMC03924.

OS Rhizobium meliloti (Sinorhizobium meliloti).

OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;

OC Rhizobiaceae; Sinorhizobium.

OC NCBI\_TaxID=382;

RT [1]

SEQUENCE FROM N.A.

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC



RC STRAIN-1021;  
 RA MEDLINE-21396507; PubMed-11481430;  
 RA Capela D., Barloy-Hubler F., Gouzy J., Bothe G., Ampe F., Batut J.,  
 RA Bolstad P., Becker A., Boutry M., Cadieu E., Drenon S., Gloux S.,  
 RA Godtke T., Goffeau A., Kahn D., Kiss E., Lelaune V., Masny D.,  
 RA Pohl T., Portetelle D., Puehler A., Purnelle B., Ramberger U.,  
 RA Renard C., Thebaud P., Vandenol M., Weidner S., Gallbert F.;  
 RT Analysis of the chromosome sequence of the legume symbiont  
 RT Sinorhizobium meliloti strain 1021.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 98:9877-9882(2001).  
 CC -1- FUNCTION: Synthesizes alpha-1,4-glucan chains using ADP-glucose.  
 CC -1- CATALYTIC ACTIVITY: ADP-glucose + ((1,4)-alpha-D-glucosyl)(n) =  
 CC ADP + ((1,4)-alpha-D-glucosyl)(n+1).  
 CC -1- PATHWAY: Glycogen biosynthesis; second step.  
 CC -1- SIMILARITY: BELONGS TO THE BACTERIAL/PLANT GLYCOGEN SYNTHASE  
 CC FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; AL591792; CAC47425.1; -  
 KM Glycogen biosynthesis; Transferase; Glycosyltransferase;  
 KM Complete proteome.  
 FT BINDING 15 ADP-GLUCOSE (BY SIMILARITY).  
 SQ SEQUENCE 480 AA; 51408 MW; 84C584F6E0564097 CRC64;

Query Match 88.2%; Score 30; DB 1; Length 480;  
 Best Local Similarity 83.3%; Pred. No. 25;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
 :|||||  
 Db 55 KPYGRF 60

RESULT 5  
 YGCI\_ECOLI STANDARD: PRT; 224 AA.  
 ID YGCI\_ECOLI  
 AC 046898;  
 DT 16-OCT-2001 (Rel. 40, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Hypothetical protein ygci precursor.  
 GN YGCI OR B2757.  
 OS Escherichia coli.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxID=562;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-K12 / MG1655;  
 RX MEDLINE-97426617; PubMed-9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V.,  
 RA Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F.,  
 RA Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Goeke D.J.,  
 RA Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12";  
 RL Science 277:1453-1474(1997).  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----

DR EMBL; U29579; AAA69267.1; ALT\_INIT.  
 DR EMBL; AE000359; AAC75799.1; ALT\_INIT.  
 DR Ecogene; EG3116; YGCI.  
 KM Hypothetical protein; Signal; Complete proteome.  
 FT SIGNAL 1  
 FT CHAIN 18 224  
 FT SIGNAL 17  
 FT CHAIN 18 224  
 SQ SEQUENCE 224 AA; 25209 MW; 1C42CC009B317D68 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 224;  
 Best Local Similarity 83.3%; Pred. No. 20;  
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
 :|||||  
 Db 25 RPYGRF 30

RESULT 6  
 THTR\_CHICK STANDARD: PRT; 289 AA.  
 ID THTR\_CHICK  
 AC P25324;  
 DT 01-MAY-1992 (Rel. 22, Created)  
 DT 01-MAY-1992 (Rel. 22, Last sequence update)  
 DT 30-MAY-2000 (Rel. 39, Last annotation update)  
 DE Thiosulfate sulfotransferase (EC 2.8.1.1) (Rhodanese).  
 GN TST.  
 OS Gallus gallus (Chicken).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;  
 OC Gallus.  
 OX NCBI\_TaxID=9031;  
 RN [1]  
 RP SEQUENCE.  
 RC TISSUE=liver;  
 RX MEDLINE-91113289; PubMed-2275748;  
 RA Kohanski R.A., Helinikson R.L.;  
 RT "Primary structure of avian hepatic rhodanese.";  
 RL J. Protein Chem. 9:369-377(1990).  
 CC -1- FUNCTION: FORMATION OF IRON-SULFUR COMPLEXES AND CYANIDE  
 CC DETOXIFICATION.  
 CC -1- CATALYTIC ACTIVITY: Thiosulfate + cyanide = sulfite + thiocyanate.  
 CC -1- SUBUNIT: MONOMER.  
 CC -1- SUBCELLULAR LOCATION: Mitochondrial matrix.  
 CC -1- TISSUE SPECIFICITY: FOUND IN NUMEROUS TISSUES.  
 CC -1- DOMAIN: THE STRUCTURE CONSISTS OF 2 DOMAINS OF VERY SIMILAR  
 CC CONFORMATION, SUGGESTING A COMMON EVOLUTIONARY ORIGIN. HOWEVER,  
 CC THE SEQUENCES OF THE 2 DOMAINS ARE VERY DIFFERENT.  
 CC -1- SIMILARITY: BELONGS TO THE RHODANESE FAMILY.  
 CC PIR; A37209; A37209.  
 DR HSSP; P00586; IRHS.  
 DR InterPro; IPR001307; Rhodanese.  
 DR InterPro; IPR001763; Rhodanese\_domain.  
 DR Pfam; PF00581; Rhodanese; 2.  
 DR SMART; SM00450; RHOD. 2.  
 DR PROSITE; PS00380; RHODANESE\_1; 1.  
 DR PROSITE; PS00683; RHODANESE\_2; 1.  
 KM Transferase; Mitochondrion.  
 FT DOMAIN 1 142 A DOMAIN.  
 FT DOMAIN 143 158 HINGE.  
 FT DOMAIN 159 289 B DOMAIN.  
 FT ACT\_SITE 186 186 MAY PLAY A ROLE IN SUBSTRATE BINDING (BY  
 FT ACT\_SITE 244 244 SIMILARITY).  
 FT ACT\_SITE 245 245 BY SIMILARITY.  
 FT ACT\_SITE 246 246 SUBSTRATE (THIOSULFATE) BINDING  
 FT ACT\_SITE 246 246 (BY SIMILARITY).  
 FT ACT\_SITE 246 246 SUBSTRATE (THIOSULFATE) BINDING  
 SQ SEQUENCE 289 AA; 32286 MW; 8BFC671DE0B2BA4 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 289;  
 Best Local Similarity 83.3%; Pred. No. 26;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 RPVGRE 6  
:|||||  
Db 182 RPAGRF 187

RESULT 7  
ID AR72\_HUMAN STANDARD; PRT; 330 AA.  
AC 043488: 075749;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Aflatoxin B1 aldehyde reductase 1 (EC 1.1.1.17) (AFB1-AR 1)  
GN Aflatoxin B1 aldehyde reductase 1 (AFB1-AR 1)  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Liver;  
RX MEDLINE=98244807; Pubmed=9576847;  
RA Ireland L.S., Harrison D.J., Neal G.E., Hayes J.D.;  
RT "Molecular cloning, expression and catalytic activity of a human AKR7  
member of the aldo-keto reductase superfamily: evidence that the  
RT major 2-carboxybenzaldehyde reductase from human liver is a homologue  
RT of rat aflatoxin B1-aldehyde reductase.";  
RT Biochem. J. 332:21-34(1998).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC TISSUE=Brain;  
RX MEDLINE=99040634; Pubmed=9823300;  
RA Prami C., Savelyeva L., Perri P., Schwab M.;  
RT "Cloning of the human aflatoxin B1-aldehyde reductase gene at 1p35-  
RT 1p36.1 in a region frequently altered in human tumor cells.";  
RT Cancer Res. 58:5014-5018(1998).  
RN [3]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases  
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING  
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOLCOHOL. COULD BE  
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC  
CC EFFECTS OF AFLATOXIN B1. ACTS AS A 2-CARBOXYBENZALDEHYDE  
CC REDUCTASE.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF026947; AAC52104.1; -;  
DR EMBL: Y16675; CAAT6347.1; -;  
DR EMBL: AL035413; CAB7232.1; -;  
DR MIM: 603418; -;  
DR InterPro: IPR001395; Aldo\_ket\_red.  
DR Pfam: PF00248; aldo\_ket\_red; 1.  
KW Oxidoreductase.  
FT ACT\_SITE 112 112 HYDROGEN-BOND DONOR (PROBABLE).  
FT CONFLICT 113 113 A -> T (IN REF. 1).  
FT CONFLICT 113 113  
SQ SEQUENCE 330 AA; 36618 MW; 3BBFB7ED0CAFA454 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 330;  
Best Local Similarity 83.3%; Pred. No. 29;

Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRE 6  
:|||||  
Db 217 QPVGRE 222

RESULT 8  
ID AR73\_HUMAN STANDARD; PRT; 331 AA.  
AC 095154: 09NUC3;  
DT 16-OCT-2001 (Rel. 40, Created)  
DT 16-OCT-2001 (Rel. 40, Last sequence update)  
DT 16-OCT-2001 (Rel. 40, Last annotation update)  
DE Aflatoxin B1 aldehyde reductase 2 (EC 1.1.1.18) (AFB1-AR 2).  
GN AKR7A3.  
OS Homo sapiens (Human).  
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.  
OX NCBI\_TaxID=9606;  
RN [1]  
RP SEQUENCE FROM N.A., AND CHARACTERIZATION.  
RC TISSUE=Liver;  
RX MEDLINE=99315412; Pubmed=10383892;  
RA Knight L.P., Primiano T., Groopman J.D., Kensler T.W., Sutter T.R.;  
RT "cDNA cloning, expression and activity of a second human aflatoxin  
RT B1-metabolizing member of the aldo-keto reductase superfamily,  
RT AKR7A3.";  
RT Carcinogenesis 20:1215-1223(1999).  
RN [2]  
RP SEQUENCE FROM N.A.  
RA Hall R.;  
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases  
CC -1- FUNCTION: CAN METABOLIZE THE AFLATOXIN B1 (AFB1) PROTEIN-BINDING  
CC DIHYDRODIOL BY FORMING NONBINDING AFB1 DIOLCOHOL. COULD BE  
CC INVOLVED IN PROTECTION OF LIVER AGAINST THE TOXIC AND CARCINOGENIC  
CC EFFECTS OF AFLATOXIN B1.  
CC -1- SUBCELLULAR LOCATION: Cytoplasmic.  
CC -1- SIMILARITY: BELONGS TO THE ALDO/KETO REDUCTASE 2 FAMILY.  
CC -----  
CC THIS SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
DR EMBL: AF040639; AAD02195.1; -;  
DR EMBL: AL035413; CAB7232.1; -;  
DR InterPro: IPR001395; Aldo\_ket\_red.  
DR Pfam: PF00248; aldo\_ket\_red; 1.  
DR PRINTS: PR00069; ALDKETREDASE.  
KW Oxidoreductase.  
FT ACT\_SITE 113 113 HYDROGEN-BOND DONOR (PROBABLE).  
FT CONFLICT 51 51 E -> D (IN REF. 1).  
FT CONFLICT 138 138 V -> M (IN REF. 1).  
FT CONFLICT 201 201 A -> ADOSEGGCSFGWTLGPGADCCFPS (IN REF.  
FT CONFLICT 201 201 2).  
FT CONFLICT 215 215 N -> D (IN REF. 1).  
FT CONFLICT 323 323 T -> A (IN REF. 1).  
SQ SEQUENCE 331 AA; 37206 MW; B9C32C33C7102AB3 CRC64;

Query Match 85.3%; Score 29; DB 1; Length 331;  
Best Local Similarity 83.3%; Pred. No. 29;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPVGRE 6  
:|||||  
Db 218 QPVGRE 223

```
RESULT 9
Y4AD_RHISN STANDARD: PRT: 292 AA.
AC P5351:
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 15-JUL-1998 (Rel. 36, Last annotation update)
DE Hypothetical 32.5 kDa protein Y4AD.
GN Y4AD.
OS Rhizobium sp. (strain NGR234).
OC Plasmid sym pNGR234e.
OC Bacteria; Proteobacteria; alpha subdivision; Rhizobiaceae group;
OC Rhizobiaceae; Rhizobium.
ON NCBI_TaxID=394;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97305956; PubMed=9163424;
RA Freiberg C.A., Fellay R., Bairoch A., Broughton W.J., Rosenthal A.,
RA Perret X.;
RT "Molecular basis of symbiosis between Rhizobium and legumes.";
RL Nature 387:394-401(1997).
CC -1- FUNCTION: PROBABLY PART OF AN OPERON Y4ABCD INVOLVED IN THE
CC SYNTHESIS OF AN ISOPRENOID COMPOUND.
CC -1- SIMILARITY: SOME, NO PHYTOENE AND SQUALENE SYNTHETASES.
CC -1- SIMILARITY: SOME, TO Y4AC.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE000064; AAB91602.1; -
DR Hypothetical protein; Plasmid; Isoprene biosynthesis.
KW SEQUENCE 292 AA; 32359 MW; 1053E5E7E3614379 CRC64;
SQ

Query Match 82.4%; Score 28; DB 1; Length 292;
Best Local Similarity 100.0%; Pred. No. 44;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVGGRF 6
DB 127 PVGRF 131

RESULT 10
MURB_LACLA STANDARD: PRT: 299 AA.
ID MURB_LACLA
AC O9CGD5:
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 01-MAR-2002 (Rel. 41, Last annotation update)
DE UDP-N-acetylglucosaminylglucosamine reductase (EC 1.1.1.158) (UDP-N-
DE acetylmutamate dehydrogenase).
GN MURB OR IL1162.
OS Lactococcus lactis (subsp. lactis) (Streptococcus lactis).
OC Bacteria; Firmicutes; Bacillus/Clostridium group; Streptococcaceae;
OC Lactococcus.
ON NCBI_TaxID=1360;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=IL1403;
RX MEDLINE=21235186; PubMed=11337471;
RA Boletín A., Wincker P., Mauer S., Jallion O., Malarme K.,
RA Weissenbach J., Ehrlich S.D., Sorokin A.;
RT "The complete genome sequence of the lactic acid bacterium Lactococcus
RT lactis ssp. lactis IL1403."
RL Genome Res. 11:731-753(2001).
CC -1- FUNCTION: CELL WALL FORMATION (BY SIMILARITY).
CC -1- CATALYTIC ACTIVITY: UDP-N-acetylmuramate + NADP(+) -> UDP-N-
```

```
CC acetyl-3-O-(1-carboxyvinyl)-D-glucosamine + NADPH.
CC -1- COFACTOR: FAD (BY SIMILARITY).
CC -1- PATHWAY: PEPTIDOGLYCAN BIOSYNTHESIS.
CC -1- SUBCELLULAR LOCATION: Cytoplasmic (Probable).
CC -1- SIMILARITY: BELONGS TO THE MURB FAMILY.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: AE006348; AAK05260.1; -
DR InterPro: IPR003170; MurB.
DR Pfam: PF02215; MurB_1.
DR Pfam: PF02873; MurB_C; 1.
KW Oxidoreductase; NADP; Flavoprotein; FAD; Cell wall; Cell division;
KW Peptidoglycan synthesis; Complete proteome.
SQ SEQUENCE 299 AA; 32098 MW; 8CE430FD281DB4FA CRC64;
```

```
Query Match 82.4%; Score 28; DB 1; Length 299;
Best Local Similarity 83.3%; Pred. No. 45;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 1 RPVGGRF 6
DB 226 RPVGGRF 231

RESULT 11
YR33_THERPE STANDARD: PRT: 300 AA.
ID YR33_THERPE
AC P13889;
DT 01-APR-1990 (Rel. 14, Created)
DT 01-APR-1990 (Rel. 14, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Hypothetical 33.4 kDa protein in ribosomal RNA operon.
OS Thermophilum pendens.
OC Archaea; Crenarchaeota; Thermoproteales; Thermofillicaceae; Thermophilum.
ON NCBI_TaxID=2269;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=HV3 / DSM 2475;
RA Kjems J., Loeffers H., Olesen T., Ingelore H., Garrett R.A.;
RT "Sequence, organisation and transcription of the ribosomal RNA operon
RT and the downstream tRNA and protein genes in the archaeobacterium
RT Thermophilum pendens."
RL Syst. Appl. Microbiol. 13:117-127(1990).
CC -1- SIMILARITY: CONTAINS 4 CBS DOMAINS.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@sib-sib.ch).
CC -----
DR EMBL: X14835; CAA32944.1; -
DR PIR: S08244; S08244.
DR InterPro: IPR000644; CBS.
DR Pfam: PF00571; CBS; 4.
DR SMART: SM00116; CBS; 4.
KW Hypothetical protein; Repeat; CBS domain.
FT DOMAIN 8 61 CBS 1.
FT DOMAIN 87 140 CBS 2.
FT DOMAIN 150 202 CBS 3.
FT DOMAIN 224 276 CBS 4.
SQ SEQUENCE 300 AA; 33437 MW; D811A313D37A4293 CRC64;
```

Query Match 82.4%; Score 28; DB 1; Length 300;  
 Best Local Similarity 100.0%; Pred. No. 45;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVGRF 6  
 |||||  
 Db 7 PVGRF 11

## RESULT 12

CATL\_BRARE STANDARD; PRT; 526 AA.  
 ID CATL\_BRARE  
 AC Q9PT92; Q918V5; 40, Created  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DE 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Catalase (EC 1.11.1.6).  
 GN CAT.  
 OS Brachydanio rerio (zebrafish) (zebra danio).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
 OC Cypriniformes; Cyprinidae; Danio.  
 OX NCBI\_TaxId=7955;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Ken C.F., Lin C.T., Wu J.L., Shaw J.F.;  
 RT "Molecular cloning of a cDNA coding for catalase from zebrafish (Danio rerio)."  
 RL Submitted (JUL-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Gerhard G.S., Kaufman E.J., Grundy M.A.;  
 RT "Molecular cloning and sequence analysis of the Danio rerio catalase gene."  
 RL Comp. Biochem. Physiol. 127:447-457(2000).  
 CC -1- FUNCTION: OCCURS IN ALMOST ALL AEROBICALLY RESPIRING ORGANISMS AND SERVES TO PROTECT CELLS FROM THE TOXIC EFFECTS OF HYDROGEN PEROXIDE.  
 CC -1- CATALYTIC ACTIVITY: 2 H(2)O(2) = O(2) + 2 H(2)O.  
 CC -1- COFACTOR: HEME GROUP AND NADP (BY SIMILARITY).  
 CC -1- SUBUNIT: HOMOTETRAMER (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Peroxisome (By similarity).  
 CC -1- SIMILARITY: BELONGS TO THE CATALASE FAMILY.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - CC the European Bioinformatics Institute. There are no restrictions on its CC use by non-profit institutions as long as its content is in no way CC modified and this statement is not removed. Usage by and for commercial CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL: AJ007505; CAB64949.1; -  
 DR EMBL: AF170069; AAF89686.1; -  
 DR HSSP: P00432; 4BLC  
 DR ZFIN: ZDB-GENE-000210-20; cat.  
 DR InterPro: IPR002226; Catalase.  
 DR Pfam: PF00199; catalase; 1.  
 DR PRINTS: PR00067; CATALASE.  
 DR Prodom: PD000510; Catalase; 1.  
 DR PROSITE: PS00437; CATALASE\_1; 1.  
 DR PROSITE: PS00438; CATALASE\_2; 1.  
 DR OXidoreductase; Peroxidase; Iron; Heme; Hydrogen peroxide;  
 KM Peroxidase; NADP.  
 FT ACT\_SITE 75 BY SIMILARITY.  
 FT ACT\_SITE 148 BY SIMILARITY.  
 FT BINDING 358 PROXIMAL HEME LIGAND (BY SIMILARITY).  
 FT CONFLICT 110 V -> A (IN REF. 2).  
 FT CONFLICT 123 P -> S (IN REF. 2).  
 FT CONFLICT 152 T -> I (IN REF. 2).  
 FT CONFLICT 161 S -> F (IN REF. 2).  
 FT CONFLICT 352 MLO -> NAA (IN REF. 2).  
 FT CONFLICT 350

FT CONFLICT 478 M -> T (IN REF. 2).  
 SQ SEQUENCE 526 AA; 59654 MW; E1120D3796522785 CRC64;  
 Best Local Similarity 100.0%; Pred. No. 79;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVGRF 6  
 |||||  
 Db 312 PVGRF 316

## RESULT 13

MDOH\_ECOLI STANDARD; PRT; 847 AA.  
 ID MDOH\_ECOLI  
 AC P3137; P77371;  
 DT 01-OCT-1993 (Rel. 27, Created)  
 DT 01-NOV-1997 (Rel. 35, Last sequence update)  
 DT 01-MAR-2002 (Rel. 41, Last annotation update)  
 DE Periplasmic glucans biosynthesis protein mdoH.  
 GN MDOH OR B1049 OR Z1684 OR ECS1427.  
 OS Escherichia coli, and  
 OS Escherichia coli O157:H7.  
 OC Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;  
 OC Escherichia.  
 OX NCBI\_TaxId=562, 83334;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12.  
 RX MEDLINE=95020533; PubMed=7934824.  
 RA Loubens I., Debarieux L., Bohin A., Lacroix J.-M., Bohin J.-P.;  
 RT "Homology between a genetic locus (mdoA) involved in the osmoregulated biosynthesis of periplasmic glucans in Escherichia coli RT and a genetic locus (hrpM) controlling pathogenicity of Pseudomonas RT syringae."  
 RL Mol. Microbiol. 10:329-340(1993).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12 / MG1655;  
 RX MEDLINE=97426617; PubMed=9278503;  
 RA Blattner F.R., Plunkett G. III, Bloch C.A., Perna N.T., Burland V., Riley M., Collado-Vides J., Glasner J.D., Rode C.K., Mayhew G.F., Gregor J., Davis N.W., Kirkpatrick H.A., Goeden M.A., Rose D.J., Mau B., Shao Y.;  
 RT "The complete genome sequence of Escherichia coli K-12.";  
 RL Science 277:12453-12474(1997).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=K12;  
 RX MEDLINE=97061202; PubMed=8905232;  
 RA Oshima T., Alba H., Baba T., Fujita K., Hayashi K., Honjo A., Ikemoto K., Inada T., Itoh T., Kajihara M., Kanai K., Kashimoto K., Kimura S., Kitagawa M., Makino K., Masuda S., Miki T., Mizubuchi K., Mori H., Motomura K., Nakamura Y., Nashimoto H., Nishio Y., Saito N., Sempel G., Seki Y., Tagami H., Takemoto K., Wada C., Yamamoto Y., Yano M., Horiuchi T.;  
 RT "A 718-kb DNA sequence of the Escherichia coli K-12 genome RT corresponding to the 12.7-28.0 min region on the linkage map.";  
 RL DNA Res. 3:137-155(1996).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
 RX MEDLINE=21074935; PubMed=11206551;  
 RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D., Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A., Posfai G., Hackett J., Klink S., Boutin A., Shao Y., Miller L., Grobdeck E.V., Davis N.W., Lim A., Dimalanta E.T., Potamous K., Apodaca J.A., Blattner F.R.;  
 RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7.";  
 RL Nature 409:529-533(2001).  
 RN [5]

SEQUENCE FROM N.A.  
 RP STRAIN-0157:H7 / RIMD 0509952;  
 RC MEDLINE-21156231; PubMed-11258796;  
 RA Hayashi T., Makino K., Ohnishi M., Kurokawa K., Ishii K., Yokoyama K.,  
 RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
 RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
 RA Kihara S., Shiba T., Hattori M., Shilagawa H.;  
 RT "Complete genome sequence of enterohemorrhagic Escherichia coli  
 RT 0157:H7 and genomic comparison with a laboratory strain K-12.";  
 RL DNA Res. 8:11-22(2001).  
 CC -1- FUNCTION: NECESSARY FOR NORMAL GLUCOSYLTRANSFERASE ACTIVITY. IT  
 CC COULD BE A GLUCOSYL TRANSFERASE OR SIMPLY A SUBUNIT OF THIS  
 CC ENZYME.  
 CC -1- PATHWAY: OPG (OSMOREGULATED PERIPLASMIC GLUCANS) BIOSYNTHESIS.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.  
 CC -1- SIMILARITY: BELONGS TO THE GLYCOSYLTRANSFERASE FAMILY 2. STRONG,  
 CC TO P. SYRINGAE HRPM  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 CC EMBL: X64197; CAA45522.1; -;  
 DR EMBL: AE000206; AAC74133.1; -;  
 DR EMBL: D90742; BAA35484.1; -;  
 DR EMBL: AE005315; AAG55795.1; -;  
 DR EMBL: AP002535; BAB34850.1; ALT\_INIT.  
 DR PIR: S35418; S35418.  
 DR Ecocore: EG11886; mdoH.  
 DR InterPro: IPR001173; Glycosyltransferase\_2.  
 DR Pfam: PF00535; Glycosyltransferase\_2; 1.  
 DR Pfam: PF00535; Glycosyltransferase\_2; 1.  
 KM Transmembrane; Inner membrane; Transferase; Glycosyltransferase;  
 KM Complete proteome.  
 FT DOMAIN 1 139 CYTOPLASMIC (POTENTIAL).  
 FT TRANSMEM 140 160 POTENTIAL.  
 FT DOMAIN 161 193 CYTOPLASMIC.  
 FT TRANSMEM 194 214 POTENTIAL.  
 FT DOMAIN 215 425 PERIPLASMIC.  
 FT TRANSMEM 426 446 POTENTIAL.  
 FT DOMAIN 447 512 CYTOPLASMIC.  
 FT TRANSMEM 513 533 POTENTIAL.  
 FT DOMAIN 534 569 PERIPLASMIC.  
 FT TRANSMEM 570 590 POTENTIAL.  
 FT DOMAIN 591 614 CYTOPLASMIC.  
 FT TRANSMEM 615 635 POTENTIAL.  
 FT DOMAIN 636 679 PERIPLASMIC.  
 FT TRANSMEM 680 700 POTENTIAL.  
 FT DOMAIN 701 745 CYTOPLASMIC.  
 FT TRANSMEM 746 766 POTENTIAL.  
 FT DOMAIN 767 847 PERIPLASMIC.  
 FT CONFLICT 19 19 A -> G (IN REF. 1).  
 FT CONFLICT 289 289 P -> L (IN REF. 1).  
 SQ SEQUENCE 847 AA; 96937 MW; 7DCAF93640180944 CRC64;

Query Match 82.4%; Score 28; DB 1; Length 847;  
 Best Local Similarity 100.0%; Pred. No. 1.36+02;  
 Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 2 PVGRF 6  
 |||||  
 DB 98 PVGRF 102

RESULT 14  
 ID ANPA\_MOUSE STANDARD; PRT; 1057 AA.  
 AC P18933;  
 DT 01-NOV-1990 (Rel. 16, Created)

01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 16-OCT-2001 (Rel. 40, Last annotation update)  
 DE Atrial natriuretic peptide receptor (ANP-A) (ANPRA) (GC-A)  
 DE (Guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide  
 DE A-type receptor).  
 GN NPRL OR NPRA.  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-C57BL/6J;  
 RX MEDLINE-90324219; PubMed-1973687;  
 RA Pandey K.N., Singh S.;  
 RT "Molecular cloning and expression of murine guanylate cyclase/atrial  
 RT natriuretic factor receptor cDNA.";  
 RL J. Biol. Chem. 265:12342-12348(1990).  
 [2]  
 RN SEQUENCE FROM N.A.  
 RP MEDLINE-95139994; PubMed-7638126;  
 RX Schoenfeld J.R., Sehl P., Quan C., Burnier J.P., Lowe D.G.;  
 RA "Agonist selectivity for three species of natriuretic peptide  
 RT receptor-A.";  
 RL Mol. Pharmacol. 47:172-180(1995).  
 CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE  
 CC CYCLASE ACTIVITY ON BINDING OF ANP.  
 CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)  
 CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE  
 CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.  
 CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE  
 CC FAMILY.  
 CC -----  
 CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sdb.ch/announce/>  
 CC or send an email to [license@sdb.ch](mailto:license@sdb.ch)).  
 CC -----  
 CC EMBL: J05504; AAA37670.1; -;  
 DR EMBL: L31932; AAA66945.1; -;  
 DR PIR: A36568; OYMSNR.  
 DR PIR: A33088; A33088.  
 DR HSP: Q02846; IAWL.  
 DR MGI: 97371; NPRL.  
 DR InterPro: IPR001170; ANF\_rcptor.  
 DR InterPro: IPR001828; ANF\_receptor.  
 DR InterPro: IPR000719; Euk\_pkinase.  
 DR Pfam: PF01094; ANF\_receptor; 1.  
 DR Pfam: PF00211; guanylate\_cyc; 1.  
 DR Pfam: PF00069; pkinase; 1.  
 DR PRINTS: PR00255; NATPEPTIDER.  
 DR SMART: SM00044; CYCC; 1.  
 DR PROSITE: PS00458; ANF\_RECEPTORS; 1.  
 DR PROSITE: PS00452; GUANYLATE\_CYCLASES\_1; 1.  
 DR PROSITE: PS00125; GUANYLATE\_CYCLASES\_2; 1.  
 DR PROSITE: PS50011; PROTEIN\_KINASE\_DOM; 1.  
 KW Receptor; Transmembrane; Glycoprotein; Phosphorylation; Lyase;  
 KW GMP synthesis; Signal.  
 FT SIGNAL 1 28  
 FT CHAIN 29 1057 ATRIAL NATRIURETIC PEPTIDE RECEPTOR A.  
 FT DOMAIN 29 469 EXTRACELLULAR (POTENTIAL).  
 FT TRANSMEM 470 490 POTENTIAL.  
 FT DOMAIN 491 1057 CYTOPLASMIC (POTENTIAL).  
 FT DOMAIN 524 801 PROTEIN KINASE-LIKE.  
 FT DOMAIN 872 1002 GUANYLATE CYCLASE.

```
FT DISULFID 88 114 BY SIMILARITY.
FT DISULFID 192 241 BY SIMILARITY.
FT DISULFID 451 451 INTERCHAIN (PROBABLE).
FT DISULFID 460 460 INTERCHAIN (PROBABLE).
FT CARBOHYD 41 41 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 208 208 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 334 334 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 375 375 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 382 382 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 423 423 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CONFLICT 3 3 G -> R (IN REF. 1).
FT CONFLICT 39 39 L -> V (IN REF. 1).
FT CONFLICT 122 122 G -> D (IN REF. 1).
FT CONFLICT 130 130 V -> L (IN REF. 1).
FT CONFLICT 285 285 R -> E (IN REF. 1).
FT CONFLICT 301 301 A -> R (IN REF. 1).
FT CONFLICT 404 404 FS -> SP (IN REF. 1).
FT CONFLICT 590 590 H -> Q (IN REF. 1).
FT CONFLICT 652 652 G -> C (IN REF. 1).
FT CONFLICT 833 833 A -> P (IN REF. 1).
FT CONFLICT 958 958 R -> G (IN REF. 1).
FT CONFLICT 1044 1044 T -> S (IN REF. 1).
FT CONFLICT 1050 1050 E -> D (IN REF. 1).
FT CONFLICT 1055 1057 TRG -> SRA (IN REF. 1).
SQ SEQUENCE 1057 AA; 119109 MW; 53A544FB2C8EF253 CRC64;
```

```
Query Match 82.4%; Score 28; DB 1; Length 1057;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 PVGRF 6
DB 120 PVGRF 124
```

```
RESULT 15
ANPA_RAT STANDARD; PRT: 1057 AA.
ID ANPA_RAT
AC P18910;
DT 01-NOV-1990 (Rel. 16, Created)
DT 01-NOV-1990 (Rel. 16, Last sequence update)
DT 16-OCT-2001 (Rel. 40, Last annotation update)
DE Atrial natriuretic peptide receptor A precursor (ANP-A) (ANPRA) (GC-A)
DE (guanylate cyclase) (EC 4.6.1.2) (NPR-A) (Atrial natriuretic peptide
DE A-type receptor).
GN NPR1.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RA MEDLINE=09143770; PubMed=2563900;
RA Chinkers M., Garbers D.L., Chang M.S., Lowe D.G., Chln H.,
RA Goeddel D.V., Schulz S.;
RA "A membrane form of guanylate cyclase is an atrial natriuretic
RT peptide receptor."
RT Nature 338:78-83(1989).
RN [2]
RP SEQUENCE FROM N.A.
RC MEDLINE=91056089; PubMed=1978722;
RA Yamaguchi M., Rutledge L.J., Garbers D.L.;
RA "The primary structure of the rat guanylyl cyclase A/atrial
RT natriuretic peptide receptor gene."
RT J. Biol. Chem. 265:20414-20420(1990).
RN [3]
RP SEQUENCE FROM N.A.
RC MEDLINE=91352095; PubMed=1679239;
RA Duda T., Goraczniak R.M., Sharma R.K.;
RA "Site-directed mutational analysis of a membrane guanylate cyclase
RT cDNA reveals the atrial natriuretic factor signaling site.";
```

```
RL Proc. Natl. Acad. Sci. U.S.A. 88:7882-7886(1991).
CC -1- FUNCTION: RECEPTOR FOR ATRIAL NATRIURETIC PEPTIDE. HAS GUANYLATE
CC CYCLASE ACTIVITY ON BINDING OF ANP.
CC -1- CATALYTIC ACTIVITY: GTP -> 3',5'-cyclic GMP + diphosphate.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- MISCELLANEOUS: THERE SEEM TO BE AT LEAST THREE ANP RECEPTORS: TWO
CC WITH GUANYLATE CYCLASE ACTIVITY (ANP-A AND ANP-B) AND ONE (ANP-C)
CC WHICH IS PROBABLY RESPONSIBLE FOR THE CLEARANCE OF ANP FROM THE
CC CIRCULATION WITHOUT A ROLE IN SIGNAL TRANSDUCTION.
CC -1- SIMILARITY: SOME SIMILARITY WITH CONSERVED REGION OF CATALYTIC
CC DOMAIN OF PROTEIN KINASES.
CC -1- SIMILARITY: BELONGS TO ADENYLYL CYCLASE CLASS-4/GUANYLYL CYCLASE
CC FAMILY.
CC -1- SIMILARITY: CONTAINS 1 PROTEIN KINASE-LIKE DOMAIN.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: X14773; CAA32881.1; -
DR EMBL: J05677; AAA41200.1; -
DR EMBL: M74535; AAA41202.1; -
DR PIR: S03348; OYRTR.
DR HSSP: Q02846; 1AHL.
DR InterPro: IPR001170; ANF_rcptor.
DR InterPro: IPR001828; ANF_receptor.
DR InterPro: IPR001054; Guanylyl_cyclase.
DR Pfam: PF01094; ANF_receptor; 1.
DR Pfam: PF00211; guanylate_cyc; 1.
DR Pfam: PF00069; pkinase; 1.
DR PRINTS: PR00255; NATPEPTIDER.
DR SMART: SM00044; CYCC; 1.
DR PROSITE: PS00458; ANF_RECEPTORS; 1.
DR PROSITE: PS00452; GUANYLATE_CYCLASES_1; 1.
DR PROSITE: PS50125; GUANYLATE_CYCLASES_2; 1.
DR PROSITE: PS50011; PROTEIN_KINASE_DOM; 1.
KW cAMP synthetase; Signal.
KW COMP. synthetase; Signal.
FT SIGNAL 1 28
FT CHAIN 29 1057
FT DOMAIN 29 469
FT TRANSMEM 470 490
FT DOMAIN 491 1057
FT DOMAIN 524 801
FT DOMAIN 872 1002
FT DISULFID 88 114
FT DISULFID 192 241
FT DISULFID 451 451
FT DISULFID 460 460
FT CARBOHYD 41 41
FT CARBOHYD 208 208
FT CARBOHYD 334 334
FT CARBOHYD 375 375
FT CARBOHYD 382 382
FT CARBOHYD 423 423
FT CONFLICT 366 366
FT CONFLICT 392 392
SQ SEQUENCE 1057 AA; 118951 MW; 9EA9ABE85AC05816 CRC64;
```

```
Query Match 82.4%; Score 28; DB 1; Length 1057;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
OY 2 PVGRF 6
DB 120 PVGRF 124
```

Search completed: September 13, 2002, 09:30:49  
Job time: 1139 sec

---

**THIS PAGE BLANK (USPTO)**



GenCore version 4.5  
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 13, 2002, 09:29:26 : Search time 311.85 Seconds  
(without alignments)  
3.883 Million cell updates/sec

Title: US-09-446-543A-73\_COPY\_15\_21

Perfect score: 34

Sequence: 1 RPYGRFX 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 562222 seqs, 172994929 residues

Total number of hits satisfying chosen parameters: 562222

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

SPTREMBL\_19:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mmc:\*
- 8: sp\_organelle:\*
- 9: sp\_phage:\*
- 10: sp\_plant:\*
- 11: sp\_protist:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_virus:\*
- 16: sp\_bacteria:\*
- 17: sp\_archaea:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	33	97.1	420	2	Q93L27 streptomyc
2	33	97.1	445	2	050531 streptomyc
3	33	97.1	511	3	059714 schizosacch
4	32	94.1	117	13	Q9W624 carassius a
5	32	94.1	529	3	Q9C1M8 streptomyc
6	30	88.2	154	2	054139 streptomyc
7	30	88.2	173	16	P74584 synecocyst
8	30	88.2	313	3	Q9U8B2 hexamita sp
9	30	88.2	369	5	096115 drosophila
10	30	88.2	501	3	Q08229 saccharomyc
11	30	88.2	559	16	Q9Kb5 bacillus ha
12	29	85.3	109	13	Q90Y59 paratichy
13	29	85.3	114	12	Q91PL7 heparitis c
14	29	85.3	118	4	Q9P195 homo sapien
15	29	85.3	170	16	Q55547 synecocyst
16	29	85.3	172	10	Q24126 nicotiana t

17	29	85.3	200	17	Q9HM02 thermoplasm
18	29	85.3	303	16	Q92PD8 rhizobium m
19	29	85.3	315	10	Q9SCQ1 arabidopsis
20	29	85.3	315	11	Q9DCM1 mus musculu
21	29	85.3	355	10	Q9FSS1 oryza sativ
22	29	85.3	362	10	Q8C996 arabidopsis
23	29	85.3	366	10	Q9M8Y5 arabidopsis
24	29	85.3	366	11	Q9D157 mus musculu
25	29	85.3	367	10	Q9SVU5 mus musculu
26	29	85.3	378	10	Q9FI48 arabidopsis
27	29	85.3	378	17	Q9HIL7 thermoplasm
28	29	85.3	379	2	P97088 cistridium
29	29	85.3	605	4	Q9BQ21 homo sapien
30	29	85.3	605	4	Q96MB9 homo sapien
31	28	82.4	106	2	Q67984 rhodococcus
32	28	82.4	141	3	Q96W56 candida alb
33	28	82.4	161	11	Q92221 mus musculu
34	28	82.4	163	4	Q9H7F5 homo sapien
35	28	82.4	163	16	Q92T10 rhizobium m
36	28	82.4	193	12	Q9QJ59 human herpe
37	28	82.4	198	5	Q9VNS8 drosophila
38	28	82.4	200	4	Q9NPN7 homo sapien
39	28	82.4	257	16	Q9HW71 pseudomonas
40	28	82.4	257	16	Q98NM6 rhizobium l
41	28	82.4	258	11	Q921C1 mus musculu
42	28	82.4	260	2	Q88068 streptomyc
43	28	82.4	266	5	Q9VUHO drosophila
44	28	82.4	294	2	Q06001 bradyrhizob
45	28	82.4	294	10	Q92UB5 arabidopsis

## ALIGNMENTS

RESULT 1

ID	Q93L27	PRELIMINARY;	PRT;	420 AA.
AC	Q93L27;			
DT	01-DEC-2001 (TREMBLrel. 19, Created)			
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-DEC-2001 (TREMBLrel. 19, Last annotation update)			
DE	CHAIN LENGTH FACTOR-LIKE PROTEIN.			
GN	AUR2B.			
OS	Streptomyces aureofaciens.			
OC	Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;			
OC	Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.			
OX	NCBI_TaxID=1894;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=CCM3239;			
RA	Koraneec J., Bistakova J., Novakova R., Homeroova D., Rezuchova B.;			
RT	"Cloning and characterization of a new polypeptide gene cluster in			
RT	Streptomyces aureofaciens CCM3239."			
RL	Submitted (May-2001) to the EMBL/GenBank/DBJ databases.			
DR	EMBL: AY033994; AAK61719.1; ..			
SQ	SEQUENCE 420 AA: 43011 MW; 3C27E22BE8C2DEA CRC64;			

Query Match 97.1%: Score 33; DB 2; Length 420;  
Best Local Similarity 100.0%: Pred. No. 24;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 RPYGRF 6  
|||||  
Db 50 RPYGRF 55

RESULT 2  
ID 050531 PRELIMINARY; PRT; 445 AA.  
AC 050531;  
DT 01-JUN-1998 (TREMBLrel. 06, Created)  
DT 01-JUN-1998 (TREMBLrel. 06, Last sequence update)

DT 01-JUN-2001 (Tremblrel. 17, last annotation update)  
DE FAD-DEPENDENT OXIDOREDUCTASE.  
GN SC9B10.24C.  
OS Streptomyces coelicolor.  
OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
OC Actinomycetales; Streptomycineae; Streptomycetaceae; Streptomyces.  
OX NCBI\_TaxID=1902;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Oliver K., Harris D.;  
RL Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA Parkhill J., Barrell B.G., Rastandream M.A.;  
RL Submitted (NOV-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN-A3(2);  
RA MEDLINE=9700351; PubMed=8843436;  
RA Redenbach M., Kleiser H.M., Denapate D., Eichner A., Cullum J.,  
RA Kinasl H., Hopwood D.A.;  
RT "A set of ordered cosmid and a detailed genetic and physical map for  
the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
RL Mol. Microbiol. 21:77-96(1996).  
DR EMBL; AL009204; CAA15814.1; -;  
DR InterPro: IPR001575; Oxid\_FAD\_bind.  
DR Pfam: PF01565; FAD\_binding\_4; 1.  
DR PROSITE: PS00862; OX2\_COVAL\_FAD; UNKNOWN\_1.  
SQ SEQUENCE 445 AA; 45067 MW; 8C0AEPB874E38C5 CRC64;

Query Match 97.1%; Score 33; DB 2; Length 445;  
Best Local Similarity 100.0%; Pred. No. 26;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||  
Db 248 RPVGRF 253

RESULT 3  
OS9714 PRELIMINARY; PRT; 511 AA.  
ID OS9714;  
AC OS9714;  
DT 01-AUG-1998 (Tremblrel. 07, Created)  
DT 01-AUG-1998 (Tremblrel. 07, last sequence update)  
DT 01-MAR-2001 (Tremblrel. 16, last annotation update)  
DE HYPOTHETICAL 57.7 KDA PROTEIN C3B8.06 IN CHROMOSOME II.  
GN SPBC388.06.  
OS Schizosaccharomyces pombe (fission yeast).  
OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
OC Schizosaccharomyces.  
OX NCBI\_TaxID=4896;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-972;  
RA Lyne M., Rajandream M.A., Barrell B.G., Beck A., Reinhardt R.,  
RA Pohl T.;  
RL Submitted (MAR-1998) to the EMBL/GenBank/DBJ databases.  
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (POTENTIAL).  
CC -1- SIMILARITY: IN C-TERMINAL TO YEAST YCR062W.  
DR EMBL; AL022244; CAA18295.1; -;  
KW Hypothetical protein; Transmembrane.  
FT TRANSMEM 33 53 POTENTIAL.  
FT TRANSMEM 59 79 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 332 352 POTENTIAL.  
FT TRANSMEM 379 399 POTENTIAL.  
FT DOMAIN 419 424 POLY-ALA.  
FT TRANSMEM 449 469 POTENTIAL.

FT TRANSMEM 483 503 POTENTIAL.  
SQ SEQUENCE 511 AA; 57703 MW; BDDDE7650B225E CRC64;

Query Match 97.1%; Score 33; DB 3; Length 511;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||  
Db 95 RPVGRF 100

RESULT 4  
OS9624 PRELIMINARY; PRT; 117 AA.  
ID OS9624;  
AC OS9624;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
DE C-RF AMIDE.  
OS Carassius auratus (Goldfish).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Ostariophysi;  
OC Cypriniformes; Cyprinidae; Carassius.  
OX NCBI\_TaxID=7957;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC TISSUE-BRAIN;  
RA Satake H., Minakata H., Fujimoto M.;  
RT Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.  
RL EMBL; AB020024; BAA76662.1; -;  
SQ SEQUENCE 117 AA; 12879 MW; D5DC4CB2038C2B0 CRC64;

Query Match 94.1%; Score 32; DB 13; Length 117;  
Best Local Similarity 83.3%; Pred. No. 11;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
|||  
Db 70 RPVGRF 75

RESULT 5  
OS91M8 PRELIMINARY; PRT; 529 AA.  
ID OS91M8;  
AC OS91M8;  
DT 01-JUN-2001 (Tremblrel. 17, Created)  
DT 01-JUN-2001 (Tremblrel. 17, last sequence update)  
DT 01-DEC-2001 (Tremblrel. 19, last annotation update)  
DE HEAT-INDUCED CATALASE.  
OS Pleurotus sajor-caju (Oyster mushroom).  
OC Eukaryota; Fungi; Basidiomycota; Hymenomycetes; Homobasidiomycetes;  
OC Agaricales; Pleurotaceae; Pleurotus.  
OX NCBI\_TaxID=50053;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC Jeong M.-J., Park S.-C.;  
RT "Cloning of the catalase gene from Pleurotus sajor-caju.";  
RL Submitted (JUL-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AF286097; AAK15159.1; -;  
DR HSSP; P15202; 1A4E.  
DR InterPro: IPR002226; Catalase.  
DR Pfam: PF00109; Catalase; 1.  
DR PRINTS; PRO0067; CATALASE.  
DR PRODOM; PD000510; Catalase; 1.  
DR PROSITE; PS00437; CATALASE\_1; 1.  
DR PROSITE; PS00438; CATALASE\_2; UNKNOWN\_1.  
SQ SEQUENCE 529 AA; 59791 MW; AA32BE6445B6AC13 CRC64;

Query Match 94.1%; Score 32; DB 3; Length 529;  
 Best Local Similarity 83.3%; Pred. No. 51;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
 :|||||  
 Db 299 RPYGRF 304

RESULT 6  
 054139 PRELIMINARY; PRT; 154 AA.  
 AC 054139;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-JUN-1998 (TREMBlrel. 06, Last sequence update)  
 DT 01-NOV-1998 (TREMBlrel. 08, Last annotation update)  
 DE HYPOTHETICAL 16.9 KDA PROTEIN.  
 GN SC2R9.14.  
 OS Streptomyces coelicolor.  
 OC Bacteria; Firmicutes; Actinobacteria; Actinobacteridae;  
 OC Actinomycetales; Streptomycinae; Streptomycetaceae; Streptomyces.  
 OX NCBI\_TaxID=1902;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Oliver K., Harris D.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN-A3(2);  
 RA Parthill J., Barrell B.G., Rajandream M.A.;  
 RL Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE-97000351; PubMed-8843436;  
 RA Redenbach M., Kleser H.M., Denaplatte D., Eichner A., Cullum J.,  
 RA Kinshl H., Hopwood D.A.;  
 RT "A set of ordered cosmids and a detailed genetic and physical map for  
 RT the 8 Mb Streptomyces coelicolor A3(2) chromosome.";  
 RL Mol. Microbiol. 21:77-96(1996).  
 DR EMBL; AL021530; CNA1648211; -;  
 KW Hypothetical protein.  
 SQ SEQUENCE 154 AA; 16884 MW; 9D9B47F36B84CD03 CRC64;

Query Match 88.2%; Score 30; DB 2; Length 154;  
 Best Local Similarity 83.3%; Pred. No. 42;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
 :|||||  
 Db 31 RPYGRF 36

RESULT 7  
 P74584 PRELIMINARY; PRT; 173 AA.  
 AC P74584;  
 DT 01-FEB-1997 (TREMBlrel. 02, Created)  
 DT 01-FEB-1997 (TREMBlrel. 02, Last sequence update)  
 DT 01-JUN-2000 (TREMBlrel. 14, Last annotation update)  
 DE HYPOTHETICAL 20.5 KDA PROTEIN.  
 GN SLR0667.  
 OS Synecocystis sp. (strain PCC 6803).  
 OC Bacteria; Cyanobacteria; Chroococcales; Synecocystis.  
 OX NCBI\_TaxID=1148;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-97061201; PubMed-8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugitara M., Sasamoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,

RA Shimpō S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synecocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 DR EMBL; D90916; BAA18691.1; -;  
 KW Hypothetical protein; Complete proteome.  
 SQ SEQUENCE 173 AA; 20509 MW; 2E2414F099C8B2F7 CRC64;

Query Match 88.2%; Score 30; DB 16; Length 173;  
 Best Local Similarity 83.3%; Pred. No. 47;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
 :|||||  
 Db 94 RPYGRF 99

RESULT 8  
 0908B2 PRELIMINARY; PRT; 313 AA.  
 AC 0908B2;  
 DT 01-MAY-2000 (TREMBlrel. 13, Created)  
 DT 01-MAY-2000 (TREMBlrel. 13, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE CARBAMATE KINASE (EC 2.7.2.2).  
 GN CBK.  
 OS Hexamita sp.  
 OC Eukaryota; Diplomonadida; Hexamitidae; Hexamitinae; Hexamita.  
 OX NCBI\_TaxID=28003;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE-20454785; PubMed-11001147;  
 RA Dimopoulos M., Bagnara A.S., Edwards M.R.;  
 RT "Characterisation and sequence analysis of a carbamate kinase gene  
 RT from the diplomonad Hexamita inflata.";  
 RL J. Eukaryot. Microbiol. 47:499-503(2000).  
 DR EMBL; AF107491; AAF08984.1; -;  
 DR HSSP; P95474; 1E19.  
 DR InterPro; IPR001048; Aakkinase.  
 DR InterPro; IPR003964; Bact\_carbamate\_kinase.  
 DR InterPro; IPR001057; Glut\_5\_kinase.  
 DR Pfam; PF00696; aakkinase.1.  
 DR PRINTS; PR00474; GI05KINASE.  
 KW Kinase; Transferase.  
 SQ SEQUENCE 313 AA; 33739 MW; 5D3AB67B7E1927B CRC64;

Query Match 88.2%; Score 30; DB 5; Length 313;  
 Best Local Similarity 83.3%; Pred. No. 85;  
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPYGRF 6  
 :|||||  
 Db 130 RPYGRF 135

RESULT 9  
 0961L5 PRELIMINARY; PRT; 369 AA.  
 AC 0961L5;  
 DT 01-DEC-2001 (TREMBlrel. 19, Created)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last sequence update)  
 DT 01-DEC-2001 (TREMBlrel. 19, Last annotation update)  
 DE GH17930P.  
 GN CG8278.  
 OS Drosophila melanogaster (Fruit fly).  
 OC Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;  
 OC Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;  
 OC Ephydroidae; Drosophilidae; Drosophila.  
 OX NCBI\_TaxID=7227;

[1]  
RN SEQUENCE FROM N.A.  
RP STRAIN-Y, CN BW SP;  
RA Stapleton M., Brokstein P., Hong L., Agbayani A., Carlson J.,  
RA Champe M., Chavez C., Dorsett V., Farfan D., Frise E., George R.,  
RA Gonzalez M., Guarrin H., Li P., Liao G., Miranda A., Mungall C.J.,  
RA Nuno J., Paelel J., Paragas V., Park S., Phouanavong S., Wan K.,  
RA Yu C., Lewis S.E., Rubin G.M., Ceiniker S.,  
RL Submitted (AUG-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL: AY015122; AAK92946.1; -  
SQ SEQUENCE 369 AA; 41753 MW; 676CE63BB03F465D CRC64;

Query Match 88.2%; Score 30; DB 5; Length 369;  
Best Local Similarity 83.3%; Pred. No. 1e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
Db 225 RPLGRF 230

RESULT 10  
008229 PRELIMINARY; PRT; 501 AA.  
AC 008229;  
DT 01-NOV-1996 (TReMBLrel. 01, Created)  
DT 01-NOV-1996 (TReMBLrel. 01, Last sequence update)  
DT 01-MAR-2001 (TReMBLrel. 16, Last annotation update)  
DE CHROMOSOME XV READING FRAME ORF YOLO70C.  
GN YOLO70C.  
OS Saccharomyces cerevisiae (Baker's yeast).  
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.  
OX NCBI\_TaxID=4932;  
RN [1]  
RP SEQUENCE FROM N.A.  
RA Alexandraki D., Katsoulou C., Tzermita M.;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
RN [2]  
RP SEQUENCE FROM N.A.  
RA MIPS;  
RL Submitted (JUL-1996) to the EMBL/GenBank/DBJ databases.  
DR EMBL: Z74812; CAA99080.1; -  
DR SGI: S0005431; YOLO70C.  
SQ SEQUENCE 501 AA; 55923 MW; 8128246D00B9C03D CRC64;

Query Match 88.2%; Score 30; DB 3; Length 501;  
Best Local Similarity 83.3%; Pred. No. 1.4e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
Db 170 KPVGRF 175

RESULT 11  
09KBE5 PRELIMINARY; PRT; 559 AA.  
AC 09KBE5;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-JUN-2001 (TReMBLrel. 17, Last annotation update)  
DE UROCOXASE (UROCANATE HYDRATASE) (EC 4.2.1.49).  
GN HUTU OR BH1983.  
OS Bacillus halodurans.  
OC Bacteria; Firmicutes; Bacillus/Clostridium group;  
OC Bacillus/Staphylococcus group; Bacillus.  
OX NCBI\_TaxID=86665;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN-C-125 / JCM 9153;

RX MEDLINE=20512582; PubMed=11058132;  
RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,  
RA Fuji F., Hiram C., Nakamura Y., Ogasawara N., Kuhara S.,  
RA Horikoshi K.;  
RT "Complete genome sequence of the alkaliphilic bacterium Bacillus  
RT halodurans and genomic sequence comparison with Bacillus subtilis."  
RL Nucleic Acids Res. 28:4317-4331(2000).  
DR EMBL: AP001513; BAB05702.1; -  
DR InterPro: IPR000193; Urocanase.  
DR Pfam: PF01175; Urocanase. 1.  
DR PRODOM: PD025423; Urocanase. 1.  
DR PROSITE: PS01233; UROCANASE; 1.  
KW Lyase; Complete proteome.  
SQ SEQUENCE 559 AA; 61603 MW; 570277A6313105A8 CRC64;

Query Match 88.2%; Score 30; DB 16; Length 559;  
Best Local Similarity 83.3%; Pred. No. 1.5e+02;  
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
Db 84 KPVGRF 89

RESULT 12  
090Y59 PRELIMINARY; PRT; 109 AA.  
AC 090Y59;  
DT 01-DEC-2001 (TReMBLrel. 19, Created)  
DT 01-DEC-2001 (TReMBLrel. 19, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE INTERLEUKINE-8.  
GN IL-8.  
OS Paralicthys olivaceus (Flounder).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;  
OC Acanthomorpha; Acanthopterygii; Percomorpha; Pleuronectiformes;  
OC Pleuronectoidae; Paralicthyidae; Paralicthys.  
OX NCBI\_TaxID=8255;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=21435923; PubMed=11551708;  
RA Lee E.Y., Park H.H., Kim Y.T., Chung J.K., Choi T.J.;  
RT "Cloning and sequence analysis of the interleukin-8 gene from flounder  
RT (Paralicthys olivaceus)."  
RL Gene 274:237-243(2001).  
DR EMBL: AF216646; AAL05442.1; -  
SQ SEQUENCE 109 AA; 12117 MW; C7CE18986C50A6ED CRC64;

Query Match 85.3%; Score 29; DB 13; Length 109;  
Best Local Similarity 66.7%; Pred. No. 50;  
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

OY 1 RPVGRF 6  
Db 42 RPIGRY 47

RESULT 13  
091PL7 PRELIMINARY; PRT; 114 AA.  
AC 091PL7;  
DT 01-OCT-2000 (TReMBLrel. 15, Created)  
DT 01-OCT-2000 (TReMBLrel. 15, Last sequence update)  
DT 01-DEC-2001 (TReMBLrel. 19, Last annotation update)  
DE GENOME POLYPROTEIN [CONTAINS: ENVELOPE GLYCOPROTEIN E2 (GP68) (GP70)  
DE (NS1)] (FRAGMENT).  
OS Hepatitis C virus.  
OC Viruses; ssRNA positive-strand viruses, no DNA stage; Flaviviridae;  
OC Hepacivirus.  
OX NCBI\_TaxID=11103;

```

RN [1]
RP SEQUENCE FROM N.A.
RA Tomaru U., Sugiyama K., Matsuno Y., Kato N.;
RT "Splenic Malignant Lymphoma with Hepatitis C Virus Infection.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
CC -1- SIMILARITY: TO HEPATITIS C VIRUS ENVELOPE GLYCOPROTEIN E1.
DR EMBL: AB044725; BAB03672.1; -.
DR InterPro: IPR002519; HCV_env.
DR InterPro: IPR002531; HCV_NSI.
DR Pfam: PF01539; HCV_env; 1.
DR Pfam: PF01560; HCV_NSI; 1.
DR ProDom: PD186062; HCV_NSI; 1.
DR Coat protein; Envelope protein; Glycoprotein; Nonstructural protein;
KW Polyprotein; Transmembrane.
FT NON_TER 1
FT SEQUENCE 114 AA; 12282 MW; 2CD020A60105F27A CRC64;
SQ
Query Match
Best Local Similarity 85.3%; Score 29; DB 12; Length 114;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;
OY 1 RPYGRF 6
DB 92 RPIGRF 97
RESULT 14
O9P195 PRELIMINARY; PRT; 118 AA.
AC O9P195;
DT 01-OCT-2000 (TREMBLrel. 15, Created)
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)
DE 01-OCT-2000 (TREMBLrel. 15, Last annotation update)
DE PRO1722.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=LIVER;
RA Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M.,
RA He F.;
RT "Functional prediction of the coding sequences of 79 new genes deduced
RT by analysis of cDNA clones from human fetal liver.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DBJ databases.
DR EMBL: AF119851; AAF69605.1; -.
SQ SEQUENCE 118 AA; 13067 MW; D6716A242C773055 CRC64;
Query Match
Best Local Similarity 85.3%; Score 29; DB 4; Length 118;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 RPYGRF 6
DB 93 QPVGRF 98
RESULT 15
O55547 PRELIMINARY; PRT; 170 AA.
AC O55547;
DT 01-NOV-1996 (TREMBLrel. 01, Created)
DT 01-NOV-1996 (TREMBLrel. 01, Last sequence update)
DT 01-DEC-2001 (TREMBLrel. 19, Last annotation update)
DE HYPOTHEICAL 18.6 KDA PROTEIN.
GN SL0293.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN-PCC 6803;
RA Tabata S.;
RL Submitted (AUG-1995) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=96127529; PubMed=8590279;
RA Kaneko T., Tanaka A., Sato S., Kotani H., Sazuka T., Miyajima N.,
RA Sugitara M., Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. I. Sequence features in the 1 Mb
RT region from map positions 64% to 92% of the genome.";
RL DNA Res. 2:153-166 (1995).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97061201; PubMed=8905231;
RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,
RA Miyajima N., Hiroseawa M., Sugitara M., Sasamoto S., Kimura T.,
RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,
RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,
RA Tabata S.;
RT "Sequence analysis of the genome of the unicellular cyanobacterium
RT Synechocystis sp. strain PCC6803. II. Sequence determination of the
RT entire genome and assignment of potential protein-coding regions.";
RL DNA Res. 3:109-136 (1996).
DR EMBL: D63999; BAA10045.1; -.
KW Hypothetical protein; Complete proteome.
SQ SEQUENCE 170 AA; 18583 MW; 6BC7097A9BB3CBF3 CRC64;

```

```

Query Match
Best Local Similarity 85.3%; Score 29; DB 16; Length 170;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
OY 1 RPYGRF 6
DB 46 QPVGRF 51

```

Search completed: September 13, 2002, 09:29:26  
Job time: 1071 sec

**THIS PAGE BLANK (USPTO)**